Universidade de Aveiro Departamento de Química
2017

Paulo André Dias Bastos

DESENVOLVIMENTO DE ENSAIOS *MULTIPLE REACTION MONITORING* PARA DIAGNÓSTICO DE CANCRO DA BEXIGA A PARTIR DE AMOSTRAS DE URINA

DEVELOPMENT OF MULTIPLE REACTION MONITORING ASSAYS FOR BLADDER CANCER DIAGNOSIS FROM URINE SAMPLES

Dissertação apresentada à Universidade de Aveiro para cumprimentos dos requisitos necessários à obtenção do grau de Mestre em Bioquímica, realizada sob a orientação da Doutora Rita Maria Pinho Ferreira, Professora Doutora do Departamento de Química da Universidade de Aveiro.

Adviser
Rita Ferreira, Department of Chemistry, University of Aveiro, Campus de Santiago, 3810-193 Aveiro, Portugal, ritaferreira@ua.pt

THE JURI

President **Ieronymos Zoidakis**

Postdoctoral Research at the Biomedical Research Foundation, Academy of Athens

Professor Antonia Vlahou

Staff Research Scientist - Professor Level at the Biomedical Research Foundation, Academy of Athens

George Mermelekas

Postdoctoral Research at the Biomedical Research Foundation, Academy of Athens

ACKNOWLEDGEMENTS

This thesis is dedicated to professors Rita Maria Pinho Ferreira and Rui Miguel Pinheiro Vitorino from University of Aveiro and to Antonia Vlahou, Ieronymos Zoidakis and George Mermelekas from Biomedical Research Foundation, Academy of Athens. Thank you for instigating my interest in science and guiding me all the way through the projects.

Resumo O Carcinoma da Bexiga é uma doença maligna com extremas implicações físicas e psicológicas para os pacientes e de elevadas repercussões socioeconómicas. A falta de procedimentos de diagnóstico precoce não-invasivos tem permitido que a sobrevivência destes pacientes tenha permanecido inalterada nos últimos 30 anos. Desta forma, biomarcadores para diagnóstico não-invasivo são urgentemente necessários, e amostras de urina representam o meio mais promissor para alcançar este fim. Contudo, apesar de várias tentativas, ensaios imunológicos realizados em amostras de urina demonstram fraca performance clínica e analítica.

Single/Multiple Reaction Monitoring (SRM/MRM) é uma técnica de espectrometria de massa para quantificação exata e absoluta. SRM/MRM representa a alternativa mais promissora para efeitos de quantificação, sendo altamente reprodutível, sensível e robusta. Desta forma, objetivou-se o desenvolvimento de ensaios por SRM/MRM para quantificação de biomarcadores de cancro da bexiga na urina, combinando múltiplos marcadores num classificador unificador.

O ensaio MRM desenvolvido demonstrou em exatidão e especificidade equiparável ou superior aos ensaios imunológicos até á data disponível. Combinando SLIT2, PROF1, SPRC e NMP22 num classificador baseado em 4 marcadores resultou em performance clínica comparável (~70% sensibilidade e ~100% especificidade ou ~80% sensibilidade e ~57% especificidade) quando comparado com os ensaios convencionais. Contudo, a quantificação livre de interferências não pode ser assegurada devido a efeitos da matriz. Um método eficiente e reprodutível para remover substâncias contaminantes presentes na urina sem comprometer a deteção dos marcadores em causa é necessária para atenuar os efeitos de matriz.

Palavras-chave *Multiple Reaction Monitoring*, Cancro da bexiga, Espectrometria de massa, Biomarcador, Urina

Abstract Bladder cancer is a malignant disease with extreme physical and psychological implications for the patients together with major economic societal costs. The lack of early non-invasive diagnostic procedures has allowed survival outcomes to remain unaltered for the past 30 years. Accordingly, non-invasive diagnostic biomarkers are urgently needed, and urine samples represent the most promising means for non-invasive bladder cancer diagnosis. However, despite several encouraging claims, available immuno-based molecular assays display poor analytical and clinical performance in urine samples. Single/Multiple Reaction Monitoring (SRM/MRM) is a high-performance mass spectrometry scanning mode for precise targeted quantification. SRM/MRM represents the most promising approach for biomarker quantification purposes, as it is highly reproducible, sensitive and robust. The main aim of this thesis was thus to develop a SRM/MRM-based assay for bladder cancer urinary biomarker quantification, combining multiple markers into a unifying classifier. In addition, two independent chapters have been dedicated to i) the value of urine proteomics for disease diagnostics and to ii) the burden of the disease together with available tools for its diagnosis in the form of a literature meta-analysis and book chapter, respectively.

At the individual biomarker level, the MRM assay herein developed for urine profiling provided comparable-to-superior accuracy and specificity as comparedwhen to ELISA assays. The combination of SLIT2, PROF1, SPRC and NMP22 in a 4-marker classifier resulted in comparable-to-superior clinical performance (~70% sensitivity with ~100% specificity or ~80% sensitivity with ~57% specificity) over conventional immuno-based assays. However, interference-free measurements still could not be assured due to urinary matrix effects. A cost-efficient and reproducible method for the removal of unidentified urinary contaminating substances without compromising the signal for the sought biomarkers is required in order to counteract urinary matrix effects.

Keywords Multiple Reaction Monitoring, Bladder Cancer, Mass spectrometry, Biomarker, Urine

Table of Contents

SECTION I – INTRODUCTORY SECTION 6
A – PAPER: *Deciphering the Disease-related Molecular Networks Using Urine Proteomics* 7
B – BOOK CHAPTER: *The Role of Urinary Proteases in Bladder Cancer* 40
C - BIOMARKER VALIDATION 81
THEORETICAL BACKGROUND 82
SECTION II – EXPERIMENTAL METHODOLOGY 101
SECTION III – EXPERIMENTAL RESULTS, DISCUSSION AND CONCLUSION 108
***Results and Discussion* 109**
Automated Peak Quality Assignments 109
Quality Control Assays 111
QC Assays: Heavy Peptides Linearity in Buffer A 111
QC Assays: Percentage Recovery and Analytical Performance 113
QC Assays: Constant amount of heavy peptides, variable amount of endogenous protein/sample matrix 116
Reproducibility 119
Clinical Samples Analysis 121
***Conclusions* 139**
***Future Directions* 140**
REFERENCES 141

SECTION I – INTRODUCTORY SECTION

A – PAPER: *Deciphering the Disease-related Molecular Networks Using Urine Proteomics*

Paulo André Dias Bastos[1], Antonia Vlahou[2], Adelino Leite-Moreira[3] , Lúcio Lara Santos[4,5], Rita Ferreira[6,*] and Rui Vitorino[1,3*]

[1]Department of Medical Sciences, Institute for Biomedicine - iBiMED, University of Aveiro, Aveiro, Portugal
[2]Biomedical Research Foundation, Academy of Athens, Athens, Greece School of Biomedical & Healthcare Sciences, Plymouth University Peninsula Schools of Medicine and Dentistry, Plymouth, UK.
[3]Department of Physiology and Cardiothoracic Surgery, Faculty of Medicine, University of Porto, Porto, Portugal
[4]Experimental Pathology and Therapeutics Group - Research Center, Portuguese Oncology Institute - Porto (IPO-Porto), Porto, Portugal
[5]Department of Surgical Oncology, Portuguese Oncology Institute - Porto (IPO-Porto), Porto, Portugal.
[6]Mass Spectrometry Group, QOPNA, Department of Chemistry, University of Aveiro, Aveiro, Portugal

[*]Corresponding author
Rui Vitorino
Department of Medical Sciences, University of Aveiro
3810-193 Aveiro
Department of Physiology and Cardiothoracic Surgery, Faculty of Medicine, University of Porto, Porto,
4000 Porto
Portugal
e-mail: rvitorino@ua.pt

NOTE: This paper has been published by the ***Trends in Analytical Chemistry*** journal.

P.A. Dias Bastos, A. Vlahou, A. Leite-Moreira, L.L. Santos, R. Ferreira, R. Vitorino, Deciphering the disease-related molecular networks using urine proteomics, Trends in Analytical Chemistry. (2017), doi: 10.1016/j.trac.2017.07.018.

ABSTRACT

Despite the large number of studies focused on the impact of diseases on urine proteome, few outputs with clinical meaning were retrieved so far. The goal of this study was to identify the biological processes modulated in urine by distinct diseases to better understand disease pathogenesis and identify urinary proteins with potential diagnosis value. We searched PubMed and SCOPUS databases for experimental papers and pooled differentially expressed proteins by disease and target organic system. Protein networks were established and functional roles revealed based on protein-protein interactions and protein altered levels.

A total of 2,572 differentially expressed proteins or peptides were pooled from 89 studies focused on 57 diseases. Data analysis highlighted inflammation as a biological process modulated by all diseases: however, specific inflammatory signatures are associated to specific diseases and/or aetiologies. Moreover, specific biological processes were identified for specific group of pathologies and unique proteins identified (e.g. APEX1 in bladder cancer, vinculin in prostate cancer, histone-lysine N-methyltransferase 2D and collagen fragments in diastolic dysfunction, and P2X purinoceptor 4 in Kawasaki disease). Some of these disease-specific proteins are originated in exosomes and resulted from the increased expression of fetal genes, possibly reflecting a phenotype reprogramming.

Integrative data analysis from urine proteomics/peptidomics reinforces the clinical potential of this body fluid for the clinical managing distinct diseases, highlighting novel putative biomarkers to be validated in the future using targeted approaches.

KEYWORDS Protein profiling, bioinformatics, cancer, cardiovascular diseases, autoimmune, inflammation.

INTRODUCTION

Proteomics-based studies offer the possibility of studying disease-specific associated proteins and their deregulation without prior knowledge thereof [1,2]. These proteins can be seen as potential biomarkers of a physiological or pathological process or the reflection of a therapeutic response [3]. However, the journey of a biomarker from the bench to the clinic is long and challenging [4]. One of the challenges relies on the choice of the biological sample, considering not only methodological but also patients' features. The screening of disease biomarkers in tissues has been a valuable tool in both clinical practice and research settings, but the invasiveness of the sampling procedure and sample complexity make this biological sample inappropriate for monitoring purposes. Therefore, biological fluids are preferred over tissue specimens whenever possible. Blood derived samples have been chosen for laboratorial studies, both clinical and research ones, because blood is in direct contact with most cells in the body and it is relatively easy to collect. However, this first feature also makes it an extremely complex fluid [5]. So, urine has become one of the most attractive diagnostic samples due to the non-invasiveness of its collection, unrestricted quantities obtainable, stability, possibility of being repeatedly sampled, and lower complexity compared to other fluids as blood plasma or serum [6].

Despite all the advances in urine proteomics, few protein targets are monitored in urine for clinical decision, being albumin considered the most reliable marker in diseases such as diabetes-related nephropathy and heart failure [7–10]. However, there is considerable hope in the technological advances observed in the last years to bring novel urine biomarkers into clinical practice [11]. Thus, this review aims to provide an integrative picture of the main molecular pathways modulated in urine by various diseases including cancer, infections and cardiovascular diseases. Data mining was used to provide novel insights on disease pathogenesis and to disclose potential urinary biomarkers.

METHODOLOGICAL ISSUES IN URINE PROTEOME ANALYSIS

A human being daily excretes less than 150 mg of proteins in urine, influenced mainly by protein concentration in blood and glomerular filtration, and partially by phenomena such as tubular reabsorption, secretion, and degradation [12,13]. About 70% of urinary proteins are of kidney origin, and only about 30% come from plasma [14]. The most abundant soluble protein in urine is Tamm–Horsfall protein (also known as uromodulin), though certain disease conditions can also greatly increase the concentration of albumin and other proteins, such as immunoglobulins [15]. Notwithstanding the informative value of these and other highly abundant urinary proteins, they might hamper the analysis of other smaller and less abundant proteins. The high salt content, mainly urea, also influences the analysis of urine proteome [16]. As result, urine samples are usually fractionated before protein profiling in order to remove salts and, eventually, the more abundant proteins to increase sensibility for smaller and less abundant proteins with potential diagnostic value (reviewed by [17,18]).

The urinary proteome may be significantly affected by normal physiologic variations, including circadian ones, and environmental factors. Sample timing and subjects' hydration status must therefore be considered in data interpretation, and 24h urine samples should be preferred to spot urine samples. However, to the best of our knowledge, only 3 studies have analyzed 24h urine samples [19–21]. Most frequently, urine collection consists of spontaneously voided samples, usually the midstream portion of second morning urines in order to avoid contaminations by urinary tract overnight residues. In the majority of proteomic studies, urine samples were collected either during multiple days (less frequently) or at multiple time points during a limited time period (e.g. 24h, pre-/peri-/post-operative) after an acute event. Urine samples obtained from indwelling or foley catheters are scarce for practical reasons, as do biopsied samples [22,23].

After collection, a protease inhibitor can be added if samples are not immediately analyzed, and/or samples should be frozen. By adding protease inhibitors, samples can be used to either proteomics or peptidomics studies, being subsequently hydrolyzed in the former case or left intact in the later one [24]. However, proteases inhibitors have been proven ineffective in increasing protein identification and reducing trypsin digestion efficacy. In most studies, urine samples are centrifuged, and if not immediately

analyzed, frozen at -80ºC [11,25]. Once proteins can precipitate during the freezing process, samples should be fully resolubilized when thawed. Sonication, addition of detergents or Tris buffers are typically preferred in the resolubilization procedure [25].

Urine samples are usually subjected to fractionation, pre-fractionation, and pre-concentration to simplify downstream analytical procedures or to make possible the analysis of subproteomes (e.g. proteins within a molecular weight range) [19,26–33]. Several methods have been used to concentrate and desalt proteins, including dialysis, protein precipitation, and lyophilization, which advantages and limitations were discussed by others [34,35]. Additional methods can be employed to remove the highly abundant proteins (e.g. albumin, uromodulin). These include targeting specific solute compartments such as urinary exosomes, [36] and affinity fractionation strategies based on immunodepletion or functionalized beads [37]. The latter technique is based on millions of random hexapeptide ligands, [38] whereas immunodepletion is based on depletion of only 6 to 10, at most, highly abundant proteins [39]. In the characterization of urine proteome/peptidome the most popular approaches are reliant on mass spectrometry (MS), which approaches are overviewed at Supplementary Table S1.

Despite advances in MS-based technology, we are now facing the problem of standardization of, for instance, methodologies, results reporting, and definition of reference values. A new subset of questions arises, such as what would be the best approach, if there is one, and what do differential expression values of a given protein mean for each person in a clinical perspective.

DATA MINING

In order to improve the diagnosis value of urine proteome, a comparative analysis of urine protein profile across several pathophysiological conditions was performed. Data research was carried out from November 2015 to January 2016. We searched PubMed and SCOPUS databases for articles published in the last 10 years, focused on differentially expressed proteins (DEPs) in urine collected from diseased and healthy human subjects. As inclusion criteria, studies had to employ mass spectrometry-based techniques and only statistically significant DEPs (whenever information regarding so

was available) were considered. At the end, data from 89 articles was pooled. For detailed information see Supplementary Information 1.

INSIGHTS ON THE FUNCTIONAL MEANING OF URINE PROTEOME

The study of urinary proteome already resulted in the identification of 4,783 distinct proteins in healthy conditions (listed at Supplementary Table S3 and retrieved from [40,41]). The characterization of exosomes contributed in 33% to urine proteome (Supplementary Table S4). Proteins differentially expressed in 57 pathophysiological conditions were pooled together from the analysis of 89 studies (Supplementary Table S2 and Table S5) which resulted in a list of 2,572 DEPs. We considered the DEPs defined in each selected paper despite not being consensual among different studies, once it was adjusted to the methodological approach used and the number of patients included in each study. Most of these studies were focused on renal diseases (23 papers), resulting in the identification of 1,054 DEPs, which represents 41% of total DEPs modulated in urine by diseases. Within renal diseases, the conditions with the largest number of DEPs identified are chronic kidney disease (273 DEPs), nephrotic syndrome (186 DEPs) and IgA nephropathy (183 DEPs). Bladder cancer has also been extensively studied, with eleven studies dedicated to it and 229 DEPs identified. Still, diseases affecting almost all organic systems, excluding the integumentary system, have been studied using urinary proteomics/peptidomics. The analysis of protein frequency across all diseases highlighted three proteins significantly modulated by most of diseases. Fibrillar forming collagen (COL1A1) was upregulated in 16 pathological conditions, collagen alpha-1 (III) (COL3A1) in 12, and alpha-1-antitrypsin in 19 diseases (Supplementary Table S5). These three proteins are known to provide strength and support to many tissues in the body, regulate cell adhesion and motility, promote cell survival and protect tissues from proteolytic degradation by inflammatory cells [42–44]. Our integrative analysis of urine proteome also showed that approximately 85% of the reported disease-related DEPs were not specific of a single disease. From the remaining, 15% were specific of a disease and 61% were part of the normal urine proteome, being either up- or downregulated. Still, 39% were exclusively detected in the urine during diseased states. Proteins uniquely modulated by a particular group of diseases (e.g.

cardiovascular diseases, infectious diseases) are presented in Supplementary Table S5 and represent the best candidate biomarkers for distinguishing diseases. This list includes 285 exclusive putative urinary biomarkers (Supplementary Table S6). In contrast to these disease-specific DEPs, proteins involved in immune response were found overrepresented in the urine of patients with distinct diseases and regardless the tissue of origin. Nevertheless, distinct inflammatory signatures characterize each unique group of diseases (e.g. cancer, infectious and autoimmune diseases), as highlighted in figure 1. For instance, it has been long established that chronic inflammation can induce tumor development, progression and metastatic dissemination, and the development of cells resistant to treatment [45]. However, inflammatory mediators can also activate antigen-presenting cells and promote antigen-mediated tumor repression, and the human host has several mechanisms to perceive and eliminate malignant cells [46]. In infectious diseases, macrophages and other phagocytes are recruited to eliminate the invading microorganism. However, this inflammatory response can be mediated by several noxious agents that also cause tissue damage [47].

To better elucidate this disease-specific immune signature, a functional analysis considering only the immune response mediators was performed. Data showed that while these mediators in cancer were primarily involved in the regulation of complement activation, those of autoimmune diseases were associated to the positive regulation of T cell mediated cytotoxicity and antigen processing and presentation of exogenous peptide antigen via MHC class I, whereas infectious diseases were found to be significantly associated to deregulations in macrophage activation, regulation of natural killer (NK) cell activation, and antigen processing and presentation of exogenous peptide antigen via MHC class I.

The role of complement in cancer is paradoxal at first sight. Our natural organic defences are known to fight cancer by promoting a pro-inflammatory state comprising increased complement activation, which is demonstrated by antibody-initiated complement-dependent cytolysis and antibody-dependent cell-mediated cytolysis [48]. Paradoxically, complement C3 activation can promote tumor growth, while complement C5a receptor signalling pathway activation enhances infiltration by myeloid-derived suppressor cells [49]. However, it is increasingly recognized that the role of the

complement system goes beyond hemostasis, with chronic low-grade activation allowing the complement system to function as a "sonar or radar system" [50,51]. Autoimmune diseases are supposed to be initiated by the activation of antigen specific T cells. Then, T cells activate self-reactive B cells and macrophages, leading to harmful autoantibody responses and direct tissue damage [52]. T cells proliferation is mediated by the recognition of self-MHC/peptide ligands. While this recognition is required for T cells survival, it leads to strong proliferation under lymphopenic conditions [53,54]. Therefore, apart from the dependence on genetic factors, previous exposure to peptide ligands and conditions leading to lymphopenia seem to be triggers of autoimmune diseases [55,56]. In infectious diseases, activated NK cells infiltrate infected tissues [57]. NK cells in lymphoid tissues are activated by dendritic cells, [58] but infected cells may upregulate the cell-surface expression of MHC class I molecules and evade NK cell-mediated killing [59]. During infection, circulating monocytes are recruited to local tissues, where they change phenotype and convert into macrophages, recognize and phagocyte pathogens, and adopt an inflammatory state that promotes tissue repair [60]. Though not surprising that distinct groups of diseases induce different inflammatory/immune responses, it is remarkable that the specificity of such inflammatory response in each of these groups of diseases can be accurately monitored by urine protein profiling.

Multiple tissues seem to contribute to urine protein profile in pathological conditions, suggesting that proteins upregulated in distinct organs can pass throughout the blood circulation and the urinary tract to take part in urine composition. In an attempt to map urinary proteins to its tissue of origin, a HeatMap was created. This comparative analysis was based on the match between identified DEPs and protein databases from human tissues obtained from either live or postmortem samples [61]. Then, those proteins that could be precisely mapped to 1 to 3 organs were manually curated and a list of approximately 100 proteins was created (Supplementary Figure S1 and Table S7). Many of these proteins are of fetal origin, which means that they are expressed at least 10-fold higher in fetal samples as compared to adult tissues and cells [61]. Several diseases seem to be accompanied by increased expression of fetal genes, most likely in an attempt to recover from tissue damage, to promote cellular growth and division, and to

adapt the organism to new requirements. Despite being normally silenced during the adult life, a genome-wide demethylation of fetal genes accounts for phenotype reprogramming under special circumstances, as those of malignant diseases [62,63]. Therefore, methylation-directed studies could be a promising avenue of research targeting the inhibition of fetal genes to avoid rapid and excessive cell proliferation.

To better explore the potential diagnosis value of urinary proteome, an integrative analysis of proteome data *per* group of diseases was performed, with emphasis in infectious diseases, cancer, autoimmune diseases, cardiovascular and renal diseases.

INFECTIOUS DISEASES

Tuberculosis, sepsis, hepatitis E virus, *Klebsiella pneumonia*, uropathogenic *Escherichia coli*, necrotizing enterocolitis, and *Histoplasma capsulatum* infections modulate the urine protein profile in a unique manner (Supplementary Table S5). For example, while the urine profile of patients infected with *K. pneumonia* is characterized by the prevalence of human proteins involved in "protein-lipid complex remodeling", "protein maturation", and "regulation of protein activation cascade", the urine of patients infected with *E. coli* is rich in proteins from "blood coagulation and fibrin clot formation", "positive regulation of substrate adhesion" and "killing of cells of other organism". Though *E. coli* is the most commonly isolated uropathogen, [64] *K. pneumonia* is an opportunistic bacteria responsible for many hospital-acquired urinary tract infections, pneumonia, septicemias and soft tissue infections [65]. Urinary proteins characteristic of *K. pneumonia* infection target erythrocyte and include hemoglobin subunits alpha and delta, band 3 anion transport protein and erythrocyte band 7 integral membrane protein (STOM) (Supplemental Table S5). Unlike *E. coli*, *K. pneumonia* is surrounded by a capsule, which increases its virulence and allows it to evade the host's immune system. As a result, infections by *K. pneumonia* may not be characterized by a cellular response as strong as the observed in infections by *E. coli* [66,67]. The capsular antigens are composed of complex acidic polysaccharides often made up of uronic acids, which limits bacteria opsonisation by complement C3, thus diminishing its phagocytosis by macrophages and contributing to the upregulation of "protein activation cascade" mediators observed in urine [67].

The bioinformatics analysis of urinary proteins modulated in human host by *E. coli* support the notion that uropathogenic *E. coli* invades the urinary tract through human host-derived factors. In this process, macrophages are involved through the activity of chitotriosidase [68] and neutrophil cytosol factors 1B, 2 and 4, which were identified in the urine of infected patients. These factors are either components of the multicomponent enzyme system NADPH-oxidase, responsible for the oxidative burst responsible for reactive oxygen species production, or involved in processing and presentation of exogenous peptide antigen via MHC class I and phagosome maturation, [68–70]. Adhesion molecules are important virulence factors contributing for *E. coli* pathogenicity [71]. While all *E. coli* strains express a wide variety of adhesion factors, uropathogenic *E. coli* isolates express variants of these factors enhancing its affinity for monomannose residues and conferring it a higher tropism towards glycoprotein receptors, extensively expressed by uroepithelial cell, thereby enhancing colonization and infections of the urinary tract [72]. Unlike the observed in infections with *K. pneumoniae*, the opsonisation of *E. coli* with complement factor C3 promotes its internalization by renal tubular epithelial cells through its interaction with the complement receptor CR1-related protein, enhancing its ability to colonize and invade the kidney [73]. Figure 2 overviews the urinary proteins uniquely modulated by these and others infectious diseases.

CANCER

Each type of cancer (prostate, uterus, bladder, colon, ovarian, and renal cell) seems to present a unique set of urinary protein markers (Supplementary Table S5). Together, these protein are involved in the “regulation of digestive system processes”, “regulation of intestinal absorption”, “response to axon injury”, “regulation of extrinsic apoptotic signalling pathway via death domain receptors”, “endothelial cell apoptotic process”, and “regulation of endothelial cell apoptotic processes” (figure 1), reflecting in part the deregulation of affected organs but also tumor-specific processes, namely alterations in apoptotic signalling pathways and neovascularization. In fact, impairment of the apoptotic signalling process is one hallmark of cancer, resulting from the accumulation of genetic defects and leading to aberrant cellular proliferation, tumorigenesis,

neovascularization and drug resistance [74,75]. Despite the low number of proteins modulated by colon cancer, when considering all proteins modulated by cancer it was noticed an overrepresentation of the "regulation of digestive system" process.

While renal cell carcinoma is characterized by a set of proteins involved mainly in "extracellular matrix disassembly", a set of proteins responsible for "protein activation cascade", "regulation of response to wounding", "protein-lipid complex remodelling", "response to fungus", and "negative regulation of immune effector" characterizes bladder cancer. In addition to the tissue remodelling caused by tumorigenesis, extracellular matrix components, especially integrins, are known to modulate cellular responses to apoptotic stimuli [76]. These extracellular matrix-dependent modulation of the apoptotic process also promotes the epithelial-mesenchymal transition, being responsible, at least in part for organ morphogenesis, tissue remodelling, embryonic development, metastasis and wound healing, [77,78] all of which were monitored by urine protein profiling. In the urine of patients with bladder cancer several apolipoproteins isoforms were found upregulated, which contributed to the overrepresentation of "protein-lipid complex remodelling" process. Still, the upregulation of apolipoproteins is not specific of bladder cancer unlike DNA-(apurinic or pyrimidinic site) lyase (APEX1), which was found exclusively upregulated in the urine of patients with bladder cancer. This enzyme functions as an apurinic/apyrimidinic (AP) endonuclease in the DNA base excision repair pathway of DNA lesions induced by oxidative and alkylating agents [79]. Its study in the set of bladder cancer has been mainly focused on risk-associated polymorphisms [80] but our data analysis suggest that high urinary levels might be seen as a marker of bladder cancer. In the set of prostate cancer, the aberrant expression of proteins like ERG and vinculin was detected through urine proteomics. ERG is a transcriptional factor overexpressed in 50% of human prostate cancers, [81] and vinculin overexpression seems to contribute to prostate cancer progression by enhancing tumour cell proliferation [82].

Figure 2 overviews the urinary proteins uniquely modulated by different types of cancer.

AUTOIMMUNE DISEASES

The biological processes more represented in the urine of patients with rheumatoid arthritis (RA), Kawasaki disease, Systemic Lupus Erythematous, Type 1 Diabetes mellitus and Transplant Rejections are “cholesterol transport”, “positive regulation of lipid transport”, “sterol transport”, “cell-cell junction organization”, and “blood coagulation” (figure 1). The increase in cholesterol transport most likely reflects an increase in lipid inflammatory mediators, which are known to trigger and mediate human diseases with an autoimmune component [83]. The upregulation of the biological process “cell-cell junction organization” is most likely due to proteins such as cell adhesion molecule 1 (CADM1) and cadherins 11 (CDH11), 13 (CDH13) and 17 (CDH17).

Looking to specific autoimmune diseases, it was noticed a high prevalence of the biological processes “protein nitrosylation”, “susceptibility to natural killer cell mediated cytotoxicity”, “positive regulation of response to biotic stimulus”, and “heterophilic cell-cell adhesion via plasma membrane cell adhesion molecules” in RA. The list of proteins involved in these biological processes is long and diverse (Supplementary Table S5); however, data analysis highlighted four proteins uniquely upregulated in the urine of RA patients, specifically complement factor H related protein 1 (CFHR1), coactosin-like protein (COLTL1), lysosomal protective protein (CTSA) and trypsin-1 (PRSS1). Despite no clear association of trypsin-1 to RA, the other three proteins seem related to the inflammatory response that characterizes this autoimmune disease. CFHR1 has a role in the complement regulation [84]. COLTL1 supports 5-lipoxygenase activity in leukotrienes, lipoxins and resolvin E1 biosynthesis, key lipid inflammatory mediators.[85] CTSA is a protective protein essential for both the activity of β-galactosidase and neuraminidase [86]. Indeed, the loss of sugars as galactose seems associated with this pathology [87].

The autoimmune diseases with more unique urinary proteins identified are Kawasaki disease and type 1 Diabetes mellitus. Among the biological processes regulated by Kawasaki disease are “antigen processing and presentation of exogenous peptide antigen via MHC class I”, “signal transduction involved in cell cycle checkpoint”, “regulation of calcium ion transport into cytosol”, “detection of external biotic stimulus”, and “regulation of protein depolymerization”. These processes were found to be mediated by 124 proteins exclusively overrepresented in Kawasaki disease

(Supplemental Table S5), as for example regulator complex protein LAMTOR1, P2X purinoceptor 4 and 60S ribosomal protein L22. LAMTOR1 is involved in cell cycle arrest, endosome localization and organization and macroautophagy,[88] and because makes part of lysosomes and late endosomes membranes, it can also be found in extracellular exosomes [89]. 60S ribosomal protein L22 is a structural constituent of ribosome responsible for alpha-beta T cell differentiation and SRP-dependent cotranslational protein targeting to membrane [90] Therefore, Kawasaki disease is significantly characterized by alterations in cell growth and cell cycle, apoptosis and proliferation-associated fundamental processes such as endothelial cell activation response to growth factors. These processes might be triggered by the inflammatory response that characterizes Kawasaki disease.

Type 1 Diabetes is characterized by proteins involved in “reverse cholesterol transport” and “triglyceride catabolic process”. Indeed, the association between type 1 Diabetes and dyslipidemia was already reported for young patients [91]. The urine protein profile of patients with type 1 Diabetes was found to be significantly different than those of patients with type 2 Diabetes (Supplementary Table S5). In fact, only 4% of the whole set of DEPs were common to both groups of patients. Overall, the main biological processes modulated by both types of Diabetes mellitus are rather unspecific, reflecting mainly the systemic alterations that characterize diabetes. However, in contrast with type1 Diabetes, type 2 Diabetes is characterized by an “acute inflammatory response” and ”deregulation in vitamin transport” and “digestive system processes”, for which contributed apolipoproteins, retinol binding protein 4 and SERPINA3.

Figure 2 overviews the urinary proteins exclusively modulated by autoimmune diseases.

CARDIOVASCULAR DISEASES

In this group of diseases that includes chronic heart failure, coronary artery disease, left ventricle diastolic dysfunction, Kawasaki disease, preeclampsia and vasculitis, the biological processes more represented in the urine were “multicellular organismal catabolism”, “bone morphogenesis” and “cellular response to amino acid stimulus”. It was noticed the upregulation of collagen synthesis, degradation and deposition, which resulted in several distinct collagen fragments. Despite being primarily correlated with

fibrosis, many of the identified collagens are also involved in cell differentiation and angiogenesis [92]. Several other proteins contributed for these biological processes, including copine-9, copine-7, osteopontin, bone morphogenetic protein (BMP) 5, and WW domain-binding protein 11. Among these, BMPs have been attributed an important role in cardiovascular diseases. Multiple BMPs-dependent events, which work in concert to attain normal cardiovascular development, control collagen remodelling and cellular proliferation, illustrate the tight connection between collagens, BMPs and the development and homeostasis of the heart [93–95]. In addition to BMPs, its receptors are also crucial for heart and vascular homeostasis. Diminished expression levels were reported in pulmonary arterial hypertension and its absence was shown to be lethal [96].

Significantly distinct biological processes were identified for two cardiovascular diseases: Kawasaki disease, which was already discussed in the setting of autoimmune diseases, and left ventricle diastolic dysfunction. In this disease a significant enrichment of the "endodermal cell differentiation", "endoderm formation", "collagen fibril organization", and "regulation of gastrulation" processes was observed. These findings highlight a high fibrotic state and the need for neovascularization in left ventricle diastolic dysfunction. The turnover of collagen fibers is mediated by fibroblasts and myofibroblasts in response to mechanical stress, vasoactive peptides and growth factors, and pro-inflammatory cytokines such as TNF-α and IL- 6 [97]. The differentiation of fibroblasts into myofibroblasts is responsible for the synthesis and degradation of collagens I and III in the heart, [97] a process that must account for the observed enrichment in differentiation and gastrulation-associated mediators in urine samples from patients with left ventricle diastolic dysfunction. Therefore, urine profiling reflects not only alterations in the balance of collagen synthesis/degradation, but also the main processes responsible for and resulting from alterations in this balance. Increased collagen synthesis over degradation results in the accumulation of collagen fibers, myocardial fibrosis, ventricular hypertrophy and diastolic dysfunction [98].

RENAL DISEASES

Most studies based on urinary proteomics and peptidomics are focused on renal and urinary track diseases (Supplemental Table S4). The comparison of proteome data among distinct renal diseases highlighted a great overlap, involving “cellular response to amino acid stimulus”, “collagen fibril organization”, “endochondral ossification”, “multicellular organism catabolism”, and “reverse cholesterol transport” (figure 1). A link between renal and cardiovascular diseases could thus be established, once both groups of diseases are significantly associated with BMPs and collagen deposition/degradation. Disorders affecting the cardiovascular or renal systems will most likely involve the other system, giving rise to a pathological condition called cardiorenal syndrome [99,100]. For instance, it is possible that BMPs involved in one of these two groups of diseases end up also affecting the other one, and this interaction may result in one disease leading to the other or both diseases developing simultaneously. Alternatively, some “protective” BMPs might be increased in order to attenuate common consequences of cardiovascular and renal diseases. Renal diseases appear to be associated with increased excretion of collagen fibers, [101] but the profile of deregulated collagens seems to differ between the two groups of diseases (Supplementary Table S5).

DISEASES OF THE DIGESTIVE SYSTEM

With the exception of Crohn’s disease, which was represented only by three unspecific proteins and was thus not amenable to bioinformatics analysis, all diseases somehow affecting organs of the digestive system were associated with unique urinary proteins (Supplementary Table S5). Hepatitis E and necrotizing enterocolitis, despite being both infectious diseases, are discussed in this section due to their target organs and clinical presentations. Hepatitis E is characterized by “retina homeostasis”, while necrotizing enterocolitis is characterized almost exclusively by “positive regulation of peptide secretion”. Proteins deregulated in Hepatitis E included membrane-associated guanylate kinase, WW and PDZ domain-containing protein 2 (MAGI2), prostaglandin-H2 D-isomerase and zinc-alpha-2-glycoprotein. MAGI2 is located at synaptosomes and is involved in cellular responses to nerve growth factor stimuli, regulating mitotic cell cycle arrest, cell migration and proliferation, neuroligin clustering for postsynaptic membrane

assembly and neuronal development [102–104]. Other PDZ proteins have already been reported as involved in viral infections. For instance, the PDZ protein PDZK1 facilitates hepatitis C virus entry into hepatocytes, resulting in hepatic fibrosis or cirrhosis, while the PDZ protein PDZD8 is known to suppress herpes simplex virus 1 replication [105,106]. In addition, human adenovirus E4-ORF1 and papillomavirus E6 proteins bind to human PDZ domains, including some members of the membrane-associated guanylate kinase (MAGUK) protein family, disrupting cellular tight junctions and promoting malignancy [107,108]. So, MAGI2 in urine might to be seen as a marker of hepatitis E infection.

The involvement of "positive regulation of peptide secretion" in necrotizing enterocolitis results from increased endopeptidase and proteolysis activities due to proteins as alpha-2-macroglobulin-like protein 1, cystatin-S, pigment epithelium-derived factor and CD14. For instance, cystatin-S is involved in the negative regulation of endopeptidase activity and proteolysis, [109] and its increase must reflect an attempt to mitigate the deleterious effects of exacerbated proteolysis.

In contrast, pancreatitis was characterized by the upregulation of "inflammatory process", "platelet degranulation", and "positive regulation of immune effector process". Proteins that distinguish pancreatitis from the remaining digestive system diseases include poliovirus receptor, which mediates NK cells adhesion with formation of a mature immunological synapse between NK and target cells. This protein triggers the secretion of lytic granules and IFN-gamma [110]. V-set and immunoglobulin domain-containing protein 4 was found increased only in the urine of patients with pancreatitis. This protein functions as a phagocytic receptor and a strong negative regulator of T-cell proliferation and IL-2 production [111].

The proteins exclusively modulated in the urine of patients with digestive diseases that are overviewed at Figure 2.

EXPLOITING THE URINARY PEPTIDOME

In order to explore the impact of diseases on urinary peptidome, differentially expressed peptides identified in several diseases were compared to a reference peptidome [112] consisting of 953 urinary peptides derived from 116 native proteins. To perform this

analysis, only those peptides with known amino acids sequence were considered, corresponding to 95 differentially expressed peptides. All differentially expressed peptides were compared with disease-related DEPs and a set of 38 unique urinary peptides that did not result from the cleavage of other proteins present in the urine was found (Supplementary Table S8). To predict the main proteases involved in the generation of these peptides Proteasix, [113] an open-source peptide-centric tool based on cleavage site specificity, was applied. Cathepsin L1 (CTSL1), kallikrein-2 (KLK2), kallikrein-6 (KLK6) and neprilysin (MME) were the proteases retrieved as the most probable ones involved in the modulation of urinary peptidome. Cathepsin L1 is a major excreted protein associated with adaptive immune response, [114] exogenous peptide antigen processing and presentation via MHC, [114] and extracellular matrix disassembly.[115] Kallikrein-2 is a glandular kallikrein that cleaves kininogen into Lys-bradykinin. This protease is primarily involved in extracellular matrix disassembly and organization, being also involved in small GTPase mediated signalling transduction [116]. Kallikrein-6 is responsible for the degradation of amyloid precursor protein, myelin basic protein, gelatin, casein, and extracellular matrix proteins, including fibronectin, laminin, vitronectin, and some collagens [117]. Neprilysin is an enzyme responsible for mediating the processing of sensory information by opioid peptides' degradation, and its upregulation might represent an organic attempt to cope the perception of pain that accompanies most diseases [118]. Taking into account the high number of urinary peptides involved in renal diseases and its repercussions in the cardiovascular system, the increased activity of neprilysin most likely reflects its involvement in the cleavage of angiotensin-1, angiotensin-2, angiotensin 1-9, and atrial natriuretic factor [119–121]. Moreover, neprilysin activity is remarkably upregulated during replicative senescence and intrinsic aging [122]. Increased activity of these four proteases is in accordance with the main biological processes modulated by human diseases in general. Despite unspecific, further studies on the activity of the peptidases responsible for urinary peptidome might provide useful insights into the molecular basis of human diseases.

CONCLUSIONS AND FUTURE PERSPECTIVES

The integrative analyse of urine proteome data highlighted the upregulation of inflammation in almost all diseases, though each disease seems to be associated with a specific inflammatory signature. Moreover, there is a set of proteins or peptides uniquely modulated by each disease and organic system involved, such as major prion protein in tuberculosis, peroxisomal acyl-coenzyme A oxidase 1 in bladder cancer, histone-lysine N-methyltransferase 2D and collagen fragments in left ventricular diastolic dysfunction, and immunoglobulin domain-containing protein 4 in pancreatitis, and P2X purinoceptor 4 in Kawasaki disease. For the enrichment of these proteins contributed exosomes, corroborating the importance of these "liquid biopsies" in mediating cellular communication and possibly conveying pathogenicity-associated factors. Another interesting finding retrieved from the integrative analysis of proteome data is the abundance in the urine of proteins copied from genes of fetal origin, revealing a deprogramming similar to the observed in early stages of development. Data analysis also highlighted cathepsin L1, kallikrein-2, kallikrein-6 and neprilysin as the proteases most probably involved in the modulation of urinary peptidome, particularly in renal and cardiovascular diseases.

Future work should target the validation of urinary proteome/peptidome data envisioning the development of multimarker panels to be implemented in the clinics for disease diagnosis and follow-up.

COMPETING INTERESTS

The authors declare that they have no competing interests.

ACKNOWLEDGEMENTS

This work was supported by the Portuguese Foundation for Science and Technology (FCT), European Union, QREN, FEDER and COMPETE for funding the iBiMED, UnIC, QOPNA research units (project UID/BIM/04501/2013, UID/IC/00051/2013, PEst-C/QUI/UI0062/2013), and COST action BM1305.

REFERENCES

[1] P. Findeisen, M. Neumaier, Mass spectrometry-based clinical proteomics profiling: current status and future directions., Expert Rev. Proteomics. 6 (2009) 457–9. doi:10.1586/epr.09.67.

[2] S.D. Patterson, R.H. Aebersold, Proteomics: the first decade and beyond., Nat. Genet. 33 Suppl (2003) 311–23. doi:10.1038/ng1106.

[3] J. Atkinson A.J., W. a. Colburn, V.G. DeGruttola, D.L. DeMets, G.J. Downing, D.F. Hoth, J. a. Oates, C.C. Peck, R.T. Schooley, B. a. Spilker, J. Woodcock, S.L. Zeger, Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework, Clin. Pharmacol. Ther. 69 (2001) 89–95. doi:10.1067/mcp.2001.113989.

[4] M.P. Pavlou, E.P. Diamandis, I.M. Blasutig, The long journey of cancer biomarkers from the bench to the clinic., Clin. Chem. 59 (2013) 147–57. doi:10.1373/clinchem.2012.184614.

[5] K. Kienzl-Wagner, J. Pratschke, G. Brandacher, Biomarker discovery in transplantation--proteomic adventure or mission impossible?, Clin. Biochem. 46 (2013) 497–505. doi:10.1016/j.clinbiochem.2012.10.010.

[6] M. Sedic, L.A. Gethings, J.P.C. Vissers, J.P. Shockcor, S. McDonald, O. Vasieva, M. Lemac, J.I. Langridge, D. Batinić, S.K. Pavelić, Label-free mass spectrometric profiling of urinary proteins and metabolites from paediatric idiopathic nephrotic syndrome., Biochem. Biophys. Res. Commun. 452 (2014) 21–6. doi:10.1016/j.bbrc.2014.08.016.

[7] E.M. Damsgaard, A. Froland, O.D. Jorgensen, C.E. Morgensen, Prognostic value of urinary albumin excretion rate and other risk factors in elderly diabetic patients and non-diabetic control subjects surviving the first 5 years after assessment, Diabetologia. 36 (1993) 1030–1036.

[8] S. Masson, R. Latini, V. Milani, L. Moretti, M.G. Rossi, E. Carbonieri, A. Frisinghelli, C. Minneci, M. Valisi, A.P. Maggioni, R. Marchioli, G. Tognoni, L. Tavazzi, Prevalence and prognostic value of elevated urinary albumin excretion in patients with chronic heart failure data from the GISSI-Heart failure trial, Circ. Hear. Fail. 3 (2010) 65–72. doi:10.1161/CIRCHEARTFAILURE.109.881805.

[9] H. Villacorta, V. Ferradaes Pde, E.T. Mesquita, A.C. Nobrega, Microalbuminuria is an independent prognostic marker in patients with chronic heart failure, Arq. Bras. Cardiol. 98 (2012) 62–69.

[10] I. De Boer, R. Katz, J. Cao, Cystatin C, albuminuria, and mortality among older adults with diabetes, Diabetes …. 32 (2009) 1833–8. doi:10.2337/dc09-0191.

[11] R.S. Lee, F. Monigatti, A.C. Briscoe, Z. Waldon, M.R. Freeman, H. Steen, Optimizing sample handling for urinary proteomics, J. Proteome Res. 7 (2008) 4022–4030. doi:10.1021/pr800301h.

[12] J.W. Froehlich, A.R. Vaezzadeh, M. Kirchner, A.C. Briscoe, O. Hofmann, W. Hide, H. Steen, R.S. Lee, An in-depth comparison of the male pediatric and adult urinary

proteomes., Biochim. Biophys. Acta. 1844 (2014) 1044–50. doi:10.1016/j.bbapap.2013.05.008.
[13] M.A.E. Valente, K. Damman, P.H.J.M. Dunselman, H.L. Hillege, A.A. Voors, Urinary Proteins in Heart Failure, Prog. Cardiovasc. Dis. 55 (2012) 44–55. doi:10.1016/j.pcad.2012.04.009.
[14] V. Thongboonkerd, Urinary proteomics : towards biomarker discovery , diagnostics and prognostics, (2008) 810–815. doi:10.1039/b802534g.
[15] I. Zubiri, F. Vivanco, G. Alvarez-Llamas, Proteomic analysis of urinary exosomes in cardiovascular and associated kidney diseases by two-dimensional electrophoresis and LC-MS/MS., Methods Mol. Biol. 1000 (2013) 209–20. doi:10.1007/978-1-62703-405-0_16.
[16] K. Rejdak, S.M. Leary, A. Petzold, A.J. Thompson, D.H. Miller, G. Giovannoni, Urinary neopterin and nitric oxide metabolites as markers of interferon beta-1a activity in primary progressive multiple sclerosis., Mult. Scler. 16 (2010) 1066–72. doi:10.1177/1352458510375100.
[17] H. Mischak, Z.A. Massy, J. Jankowski, Proteomics in uremia and renal disease., Semin. Dial. 22 (2009) 409–16. doi:10.1111/j.1525-139X.2009.00591.x.
[18] H. Mischak, W. Kolch, M. Aivaliotis, D. Bouyssié, M. Court, H. Dihazi, G.H. Dihazi, J. Franke, J. Garin, A. Gonzalez de Peredo, A. Iphöfer, L. Jänsch, C. Lacroix, M. Makridakis, C. Masselon, J. Metzger, B. Monsarrat, M. Mrug, M. Norling, J. Novak, A. Pich, A. Pitt, E. Bongcam-Rudloff, J. Siwy, H. Suzuki, V. Thongboonkerd, L.-S. Wang, J. Zoidakis, P. Zürbig, J.P. Schanstra, A. Vlahou, Comprehensive human urine standards for comparability and standardization in clinical proteome analysis., Proteomics. Clin. Appl. 4 (2010) 464–78. doi:10.1002/prca.200900189.
[19] R. Roy, G. Louis, K.R. Loughlin, D. Wiederschain, S.M. Kilroy, C.C. Lamb, D. Zurakowski, M. a. Moses, Tumor-specific urinary matrix metalloproteinase fingerprinting: Identification of high molecular weight urinary matrix metalloproteinase species, Clin. Cancer Res. 14 (2008) 6610–6617. doi:10.1158/1078-0432.CCR-08-1136.
[20] N.C. Garbett, M.L. Merchant, C.W. Helm, A.B. Jenson, J.B. Klein, J.B. Chaires, Detection of cervical cancer biomarker patterns in blood plasma and urine by differential scanning calorimetry and mass spectrometry., PLoS One. 9 (2014) e84710. doi:10.1371/journal.pone.0084710.
[21] M. Frantzi, J. Metzger, R.E. Banks, H. Husi, J. Klein, M. Dakna, W. Mullen, J.J. Cartledge, J.P. Schanstra, K. Brand, M. a. Kuczyk, H. Mischak, A. Vlahou, D. Theodorescu, A.S. Merseburger, Discovery and validation of urinary biomarkers for detection of renal cell carcinoma, J. Proteomics. 98 (2014) 44–58. doi:10.1016/j.jprot.2013.12.010.
[22] J. Metzger, T. Kirsch, E. Schiffer, P. Ulger, E. Mentes, K. Brand, E.M. Weissinger, M. Haubitz, H. Mischak, S. Herget-Rosenthal, Urinary excretion of twenty peptides forms an early and accurate diagnostic pattern of acute kidney injury., Kidney Int. 78 (2010) 1252–1262. doi:10.1038/ki.2010.322.

[23] J. Ho, M. Lucy, O. Krokhin, K. Hayglass, E. Pascoe, G. Darroch, D. Rush, P. Nickerson, C. Rigatto, M. Reslerova, Mass spectrometry-based proteomic analysis of urine in acute kidney injury following cardiopulmonary bypass: a nested case-control study., Am. J. Kidney Dis. 53 (2009) 584–95. doi:10.1053/j.ajkd.2008.10.037.

[24] L. Tsiatsiani, A.J.R. Heck, Proteomics beyond trypsin., FEBS J. 282 (2015) 2612–26. doi:10.1111/febs.13287.

[25] C. Eric Thomas, W. Sexton, K. Benson, R. Sutphen, J. Koomen, Urine collection and processing for protein biomarker discovery and quantification, Cancer Epidemiol. Biomarkers Prev. 19 (2010) 953–959. doi:10.1158/1055-9965.EPI-10-0069.

[26] Y. Wang, J. Chen, L. Chen, P. Zheng, H.-B. Xu, J. Lu, J. Zhong, Y. Lei, C. Zhou, Q. Ma, Y. Li, P. Xie, Urinary peptidomics identifies potential biomarkers for major depressive disorder., Psychiatry Res. 217 (2014) 25–33. doi:10.1016/j.psychres.2014.02.029.

[27] X.B. Ling, K. Lau, J.T. Kanegaye, Z. Pan, S. Peng, J. Ji, G. Liu, Y. Sato, T.T.S. Yu, J.C. Whitin, J. Schilling, J.C. Burns, H.J. Cohen, A diagnostic algorithm combining clinical and molecular data distinguishes Kawasaki disease from other febrile illnesses., BMC Med. 9 (2011) 130. doi:10.1186/1741-7015-9-130.

[28] A. Kentsis, Y.Y. Lin, K. Kurek, M. Calicchio, Y.Y. Wang, F. Monigatti, F. Campagne, R. Lee, B. Horwitz, H. Steen, R. Bachur, Discovery and validation of urine markers of acute pediatric appendicitis using high-accuracy mass spectrometry., Ann. Emerg. Med. 55 (2010) 62–70.e4. doi:10.1016/j.annemergmed.2009.04.020.

[29] L. Su, L. Cao, R. Zhou, Z. Jiang, K. Xiao, W. Kong, H. Wang, J. Deng, B. Wen, F. Tan, Y. Zhang, L. Xie, Identification of novel biomarkers for sepsis prognosis via urinary proteomic analysis using iTRAQ labeling and 2D-LC-MS/MS., PLoS One. 8 (2013) e54237. doi:10.1371/journal.pone.0054237.

[30] L. Su, R. Zhou, C. Liu, B. Wen, K. Xiao, W. Kong, F. Tan, Y. Huang, L. Cao, L. Xie, Urinary proteomics analysis for sepsis biomarkers with iTRAQ labeling and two-dimensional liquid chromatography-tandem mass spectrometry., J. Trauma Acute Care Surg. 74 (2013) 940–5. doi:10.1097/TA.0b013e31828272c5.

[31] G. Chen, Y. Zhang, X. Jin, L. Zhang, Y. Zhou, J. Niu, J. Chen, Y. Gu, Urinary proteomics analysis for renal injury in hypertensive disorders of pregnancy with iTRAQ labeling and LC-MS/MS, Proteomics - Clin. Appl. 5 (2011) 300–310. doi:10.1002/prca.201000100.

[32] T.K. Sigdel, A. Kaushal, M. Gritsenko, A.D. Norbeck, W.-J. Qian, W. Xiao, D.G. Camp, R.D. Smith, M.M. Sarwal, Shotgun proteomics identifies proteins specific for acute renal transplant rejection., Proteomics. Clin. Appl. 4 (2010) 32–47. doi:10.1002/prca.200900124.

[33] C.I. a Balog, O. a Mayboroda, M. Wuhrer, C.H. Hokke, A.M. Deelder, P.J. Hensbergen, Mass spectrometric identification of aberrantly glycosylated human

apolipoprotein C-III peptides in urine from Schistosoma mansoni-infected individuals., Mol. Cell. Proteomics. 9 (2010) 667–81. doi:10.1074/mcp.M900537-MCP200.

[34] A.L. Petri, C. Høgdall, I.J. Christensen, A.H. Simonsen, D. T'Jampens, M.L. Hellmann, S.K. Kjaer, E.T. Fung, E. Høgdall, Sample handling for mass spectrometric proteomic investigations of human urine, Proteomics - Clin. Appl. 2 (2008) 1184–1193. doi:10.1002/prca.200780010.

[35] P. Olszowy, B. Buszewski, Urine Sample Preparation for Proteomic Analysis., J. Sep. Sci. (2014) 1–26. doi:10.1002/jssc.201400331.

[36] M. Nawaz, G. Camussi, H. Valadi, I. Nazarenko, K. Ekstrom, X. Wang, S. Principe, N. Shah, N.M. Ashraf, F. Fatima, L. Neder, T. Kislinger, The emerging role of extracellular vesicles as biomarkers for urogenital cancers, Nat Rev Urol. 11 (2014) 688–701. doi:10.1038/nrurol.2014.301.

[37] S. Filip, K. Vougas, J. Zoidakis, A. Latosinska, W. Mullen, G. Spasovski, H. Mischak, A. Vlahou, J. Jankowski, Comparison of Depletion Strategies for the Enrichment of Low-Abundance Proteins in Urine., PLoS One. 10 (2015) e0133773. doi:10.1371/journal.pone.0133773.

[38] L. Sennels, M. Salek, L. Lomas, E. Boschetti, P.G. Righetti, J. Rappsilber, Proteomic analysis of human blood serum using peptide library beads, J. Proteome Res. 6 (2007) 4055–4062. doi:10.1021/pr070339l.

[39] W.G. Fisher, J.E. Lucas, U.F. Mehdi, D.W. Qunibi, H.R. Garner, K.P. Rosenblatt, R.D. Toto, A method for isolation and identification of urinary biomarkers in patients with diabetic nephropathy, Proteomics - Clin. Appl. 5 (2011) 603–612. doi:10.1002/prca.201000156.

[40] J. Adachi, C. Kumar, Y. Zhang, J. V Olsen, M. Mann, The human urinary proteome contains more than 1500 proteins , including a large proportion of membrane proteins, 7 (2006). doi:10.1186/gb-2006-7-9-r80.

[41] L. Santucci, G. Candiano, A. Petretto, M. Bruschi, C. Lavarello, E. Inglese, P. Giorgio Righetti, G. Marco Ghiggeri, From hundreds to thousands: Widening the normal human Urinome, Data Br. 1 (2014) 25–28. doi:10.1016/j.dib.2014.08.006.

[42] P.G.W. Gettins, Serpin structure, mechanism, and function, Chem. Rev. 102 (2002) 4751–4803. doi:10.1021/cr010170+.

[43] A.J. Ridley, Rho family proteins: Coordinating cell responses, Trends Cell Biol. 11 (2001) 471–477. doi:10.1016/S0962-8924(01)02153-5.

[44] M. Symons, J. Settleman, Rho family GTPases: More than simple switches, Trends Cell Biol. 10 (2000) 415–419. doi:10.1016/S0962-8924(00)01832-8.

[45] S. Shalapour, M. Karin, Immunity, inflammation, and cancer: an eternal fight between good and evil., J. Clin. Invest. 125 (2015) 3347–55. doi:10.1172/JCI80007.

[46] M.J. Smyth, G.P. Dunn, R.D. Schreiber, Cancer Immunosurveillance and Immunoediting: The Roles of Immunity in Suppressing Tumor Development and Shaping

Tumor Immunogenicity, Adv. Immunol. 90 (2006) 1–50. doi:10.1016/S0065-2776(06)90001-7.
[47] C. Bernardo, M.C. Cunha, J.H. Santos, J.M.C. Costa, P.J. Brindley, C. Lopes, F. Amado, R. Ferreira, R. Vitorino, L.L. Santos, Insight into the molecular basis of Schistosoma haematobium-induced bladder cancer through urine proteomics, Tumor Biol. (2016) 1–9. doi:10.1007/s13277-016-4997-y.
[48] M. Dechant, W. Weisner, S. Berger, M. Peipp, T. Beyer, T. Schneider-Merck, J.J. Lammerts Van Bueren, W.K. Bleeker, P.W.H.I. Parren, J.G.J. Van De Winkel, T. Valerius, Complement-dependent tumor cell lysis triggered by combinations of epidermal growth factor receptor antibodies, Cancer Res. 68 (2008) 4998–5003. doi:10.1158/0008-5472.CAN-07-6226.
[49] M.M. Markiewski, R.A. DeAngelis, F. Benencia, S.K. Ricklin-Lichtsteiner, A. Koutoulaki, C. Gerard, G. Coukos, J.D. Lambris, Modulation of the antitumor immune response by complement., Nat. Immunol. 9 (2008) 1225–35. doi:10.1038/ni.1655.
[50] X. Wu, D. Spitzer, D. Mao, S.L. Peng, H. Molina, J.P. Atkinson, Membrane Protein Crry Maintains Homeostasis of the Complement System, J. Immunol. 181 (2008) 2732–2740. doi:10.4049/jimmunol.181.4.2732.
[51] a P. Manderson, M.C. Pickering, M. Botto, M.J. Walport, C.R. Parish, Continual low-level activation of the classical complement pathway., J. Exp. Med. 194 (2001) 747–56. doi:10.1084/jem.194.6.747.
[52] J. Charles A Janeway, P. Travers, M. Walport, M.J. Shlomchik, Autoimmune responses are directed against self antigens, (2001).
[53] C. Viret, F.S. Wong, C.A. Janeway, Designing and maintaining the mature TCR repertoire: The continuum of self-peptide:self-MHC complex recognition, Immunity. 10 (1999) 559–568. doi:10.1016/S1074-7613(00)80055-2.
[54] Q. Ge, V.P. Rao, B.K. Cho, H.N. Eisen, J. Chen, Dependence of lymphopenia-induced T cell proliferation on the abundance of peptide/ MHC epitopes and strength of their interaction with T cell receptors., Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 1728–33. doi:10.1073/pnas.98.4.1728.
[55] J.J. Goronzy, C.M. Weyand, Thymic function and peripheral T-cell homeostasis in rheumatoid arthritis, Trends Immunol. 22 (2001) 251–255. doi:10.1016/S1471-4906(00)01841-X.
[56] S.S. Morse, N. Sakaguchi, S. Sakaguchi, Virus and autoimmunity: induction of autoimmune disease in mice by mouse T lymphotropic virus (MTLV) destroying CD4+ T cells., J. Immunol. 162 (1999) 5309–16.
[57] R. Glas, L. Franksson, C. Une, M.L. Eloranta, C. Ohlén, A. Orn, K. Kärre, Recruitment and activation of natural killer (NK) cells in vivo determined by the target cell phenotype. An adaptive component of NK cell-mediated responses., J. Exp. Med. 191 (2000) 129–38.

[58] G. Ferlazzo, M. Pack, D. Thomas, C. Paludan, D. Schmid, T. Strowig, G. Bougras, W. a Muller, L. Moretta, C. Münz, Distinct roles of IL-12 and IL-15 in human natural killer cell activation by dendritic cells from secondary lymphoid organs., Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 16606–16611. doi:10.1073/pnas.0407522101.

[59] J. Ye, B. Zhu, Z.F. Fu, H. Chen, S. Cao, Immune evasion strategies of flaviviruses, Vaccine. 31 (2013) 461–471. doi:10.1016/j.vaccine.2012.11.015.

[60] L. Zhang, C.C. Wang, Inflammatory response of macrophages in infection, Hepatobiliary Pancreat. Dis. Int. 13 (2014) 138–152. doi:10.1016/s1499-3872(14)60024-2.

[61] M.-S. Kim, S.M. Pinto, D. Getnet, R.S. Nirujogi, S.S. Manda, R. Chaerkady, A.K. Madugundu, D.S. Kelkar, R. Isserlin, S. Jain, J.K. Thomas, B. Muthusamy, P. Leal-Rojas, P. Kumar, N. a Sahasrabuddhe, L. Balakrishnan, J. Advani, B. George, S. Renuse, L.D.N. Selvan, A.H. Patil, V. Nanjappa, A. Radhakrishnan, S. Prasad, T. Subbannayya, R. Raju, M. Kumar, S.K. Sreenivasamurthy, A. Marimuthu, G.J. Sathe, S. Chavan, K.K. Datta, Y. Subbannayya, A. Sahu, S.D. Yelamanchi, S. Jayaram, P. Rajagopalan, J. Sharma, K.R. Murthy, N. Syed, R. Goel, A. a Khan, S. Ahmad, G. Dey, K. Mudgal, A. Chatterjee, T.-C. Huang, J. Zhong, X. Wu, P.G. Shaw, D. Freed, M.S. Zahari, K.K. Mukherjee, S.K.S. Shankar, A. Mahadevan, H. Lam, C.J. Mitchell, S.K.S. Shankar, P. Satishchandra, J.T. Schroeder, R. Sirdeshmukh, A. Maitra, S.D. Leach, C.G. Drake, M.K. Halushka, T.S.K. Prasad, R.H. Hruban, C.L. Kerr, G.D. Bader, C. a Iacobuzio-Donahue, H. Gowda, A. Pandey, A draft map of the human proteome., Nature. 509 (2014) 575–81. doi:10.1038/nature13302.

[62] M. Monk, M. Boubelik, S. Lehnert, Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development, Development. 99 (1987) 371–382.

[63] M. Monk, Of microbes, mice and man, Int. J. Dev. Biol. 45 (2001) 497–507.

[64] B. Foxman, The epidemiology of urinary tract infection., Nat. Rev. Urol. 7 (2010) 653–60. doi:10.1038/nrurol.2010.190.

[65] R. Podschun, U. Ullmann, Klebsiella spp. as nosocomial pathogens: Epidemiology, taxonomy, typing methods, and pathogenicity factors, Clin. Microbiol. Rev. 11 (1998) 589–603. doi:0893-8512/98/$04.00?0.

[66] I. Ofek, A. Mesika, M. Kalina, Y. Keisari, R. Podschun, H. Sahly, D. Chang, D. McGregor, E. Crouch, Surfactant protein D enhances phagocytosis and killing of unencapsulated phase variants of Klebsiella pneumoniae, Infect. Immun. 69 (2001) 24–33. doi:10.1128/IAI.69.1.24-33.2001.

[67] G. Cortes, Molecular Analysis of the Contribution of the Capsular Polysaccharide and the Lipopolysaccharide O Side Chain to the Virulence of Klebsiella pneumoniae in a Murine Model of Pneumonia, Infect. Immun. 70 (2002) 2583–2590. doi:10.1128/IAI.70.5.2583-2590.2002.

[68] R.G. Boot, G.H. Renkema, A. Strijland, A.J. Van Zonneveld, J.M.F.G. Aerts, Cloning of a cDNA encoding chitotriosidase, a human chitinase produced by macrophages, J. Biol. Chem. 270 (1995) 26252–26256. doi:10.1074/jbc.270.44.26252.
[69] T.L. Leto, K.J. Lomax, B.D. Volpp, H. Nunoi, J.M. Sechler, W.M. Nauseef, R. a Clark, J.I. Gallin, H.L. Malech, Cloning of a 67-kD neutrophil oxidase factor with similarity to a noncatalytic region of p60c-src., Science. 248 (1990) 727–30.
[70] F.B. Wientjes, J.J. Hsuan, N.F. Totty, A.W. Segal, p40phox, a third cytosolic component of the activation complex of the NADPH oxidase to contain src homology 3 domains., Biochem. J. 296 (Pt 3 (1993) 557–61.
[71] J.R. Johnson, Virulence factors in Escherichia coli urinary tract infection., Clin. Microbiol. Rev. 4 (1991) 80–128. doi:10.1128/CMR.4.1.80.Updated.
[72] E. V Sokurenko, V. Chesnokova, D.E. Dykhuizen, I. Ofek, X.R. Wu, K. a Krogfelt, C. Struve, M. a Schembri, D.L. Hasty, Pathogenic adaptation of Escherichia coli by natural variation of the FimH adhesin., Proc. Natl. Acad. Sci. U. S. A. 95 (1998) 8922–6. doi:10.1073/pnas.95.15.8922.
[73] T. Springall, N.S. Sheerin, K. Abe, V.M. Holers, H. Wan, S.H. Sacks, Epithelial secretion of C3 promotes colonization of the upper urinary tract by Escherichia coli., Nat. Med. 7 (2001) 801–6. doi:10.1038/89923.
[74] S. Negrini, V.G. Gorgoulis, T.D. Halazonetis, Genomic instability--an evolving hallmark of cancer., Nat. Rev. Mol. Cell Biol. 11 (2010) 220–228. doi:10.1038/nrm2858.
[75] J. Plati, O. Bucur, R. Khosravi-Far, Apoptotic cell signaling in cancer progression and therapy, Integr. Biol. 3 (2011) 279–296. doi:10.1039/c0ib00144a.
[76] Z. Werb, C.J. Sympson, C.M. Alexander, N. Thomasset, L.R. Lund, a MacAuley, J. Ashkenas, M.J. Bissell, Extracellular matrix remodeling and the regulation of epithelial-stromal interactions during differentiation and involution., Kidney Int. Suppl. 54 (1996) S68–S74.
[77] C. Scheel, R.A. Weinberg, Cancer stem cells and epithelial-mesenchymal transition: Concepts and molecular links, Semin. Cancer Biol. 22 (2012) 396–403. doi:10.1016/j.semcancer.2012.04.001.
[78] E. a Turley, M. Veiseh, D.C. Radisky, M.J. Bissell, Mechanisms of disease: epithelial-mesenchymal transition--does cellular plasticity fuel neoplastic progression?, Nat. Clin. Pract. Oncol. 5 (2008) 280–90. doi:10.1038/ncponc1089.
[79] S.C. Sak, P. Harnden, C.F. Johnston, A.B. Paul, A.E. Kiltie, APE1 and XRCC1 protein expression levels predict cancer-specific survival following radical radiotherapy in bladder cancer., Clin. Cancer Res. 11 (2005) 6205–11. doi:10.1158/1078-0432.CCR-05-0045.
[80] C. Liu, Q. Yin, L. Li, Y.Z. Zhuang, X. Zu, Y. Wang, APE1 Asp148Glu gene polymorphism and bladder cancer risk: a meta-analysis, Mol Biol Rep. 40 (2013) 171–176. doi:10.1007/s11033-012-2046-5.

[81] L. Wu, J.C. Zhao, J. Kim, H.-J. Jin, C.-Y. Wang, J. Yu, ERG is a critical regulator of Wnt/LEF1 signaling in prostate cancer, Cancer Res. 73 (2013) 6068–6079. doi:10.1158/0008-5472.CAN-13-0882.

[82] C. Ruiz, D.R. Holz, M. Oeggerli, S. Schneider, I.M. Gonzales, J.M. Kiefer, T. Zellweger, A. Bachmann, P.A. Koivisto, H.J. Helin, S. Mousses, M.T. Barrett, D.O. Azorsa, L. Bubendorf, Amplification and overexpression of vinculin are associated with increased tumour cell proliferation and progression in advanced prostate cancer, J. Pathol. 223 (2011) 543–552. doi:10.1002/path.2828.

[83] S. Omoigui, Cholesterol synthesis is the trigger and isoprenoid dependent interleukin-6 mediated inflammation is the common causative factor and therapeutic target for atherosclerotic vascular disease and age-related disorders including osteoporosis and type 2 diabet, Med. Hypotheses. 65 (2005) 559–69. doi:10.1016/j.mehy.2005.03.012.

[84] C. Skerka, P.F. Zipfel, Complement factor H related proteins in immune diseases, Vaccine. 26 (2008). doi:10.1016/j.vaccine.2008.11.021.

[85] D. Basavarajappa, M. Wan, A. Lukic, D. Steinhilber, B. Samuelsson, O. Rådmark, Roles of coactosin-like protein (CLP) and 5-lipoxygenase-activating protein (FLAP) in cellular leukotriene biosynthesis, Proc. Natl. Acad. Sci. . 111 (2014) 11371–11376. doi:10.1073/pnas.1410983111.

[86] N.J. Galjart, H. Morreau, R. Willemsen, N. Gillemans, E.J. Bonten, A. d'Azzo, Human lysosomal protective protein has cathepsin A-like activity distinct from its protective function, J. Biol. Chem. 266 (1991) 14754–14762.

[87] A.J. Soltys, F.C. Hay, A. Bond, J.S. Axford, M.G. Jones, I. Randen, K.M. Thompson, J.B. Natvig, The binding of synovial tissue-derived human monoclonal immunoglobulin M rheumatoid factor to immunoglobulin G preparations of differing galactose content, Scand. J. Immunol. 40 (1994) 135–143. doi:10.1111/j.1365-3083.1994.tb03442.x.

[88] Y. Sancak, L. Bar-Peled, R. Zoncu, A.L. Markhard, S. Nada, D.M. Sabatini, Ragulator-rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids, Cell. 141 (2010) 290–303. doi:10.1016/j.cell.2010.02.024.

[89] M. Prunotto, A. Farina, L. Lane, A. Pernin, J. Schifferli, D.F. Hochstrasser, P. Lescuyer, S. Moll, Proteomic analysis of podocyte exosome-enriched fraction from normal human urine, J. Proteomics. 82 (2013) 193–229. doi:10.1016/j.jprot.2013.01.012.

[90] T.I. Odintsova, E.C. Müller, A. V. Ivanov, T.A. Egorov, R. Bienert, S.N. Vladimirov, S. Kostka, A. Otto, B. Wittmann-Liebold, G.G. Karpova, Characterization and analysis of posttranslational modifications of the human large cytoplasmic ribosomal subunit proteins by mass spectrometry and Edman sequencing, J. Protein Chem. 22 (2003) 249–258. doi:10.1023/A:1025068419698.

[91] H.M. Mona, S.A. Sahar, S.M. Hend, A.-W.A. Nanees, Dyslipidemia in type 1 diabetes mellitus: Relation to diabetes duration, glycemic control, body habitus, dietary

intake and other epidemiological risk factors, Egypt. Pediatr. Assoc. Gaz. 63 (2015) 63–68. doi:10.1016/j.epag.2015.03.001.
[92] B. López, A. González, J. Díez, Circulating biomarkers of collagen metabolism in cardiac diseases., Circulation. 121 (2010) 1645–54. doi:10.1161/CIRCULATIONAHA.109.912774.
[93] K. Jiao, H. Kulessa, K. Tompkins, Y. Zhou, L. Batts, H.S. Baldwin, B.L.M. Hogan, An essential role of Bmp4 in the atrioventricular septation of the mouse heart, Genes Dev. 17 (2003) 2362–2367. doi:10.1101/gad.1124803.
[94] R.Y. Kim, E.J. Robertson, M.J. Solloway, Bmp6 and Bmp7 Are Required for Cushion Formation and Septation in the Developing Mouse Heart, Dev. Biol. 235 (2001) 449–466. doi:10.1006/dbio.2001.0284.
[95] J. Rivera-Feliciano, C.J. Tabin, Bmp2 instructs cardiac progenitors to form the heart-valve-inducing field, Dev. Biol. 295 (2006) 580–588. doi:10.1016/j.ydbio.2006.03.043.
[96] N. El-Bizri, C. Guignabert, L. Wang, A. Cheng, K. Stankunas, C.-P. Chang, Y. Mishina, M. Rabinovitch, SM22alpha-targeted deletion of bone morphogenetic protein receptor 1A in mice impairs cardiac and vascular development, and influences organogenesis., Development. 135 (2008) 2981–2991. doi:10.1242/dev.017863.
[97] T.A. Wynn, Cellular and molecular mechanisms of fibrosis., J. Pathol. 214 (2008) 199–210. doi:10.1002/path.2277.
[98] G.L. Brower, J.D. Gardner, M.F. Forman, D.B. Murray, T. Voloshenyuk, S.P. Levick, J.S. Janicki, The relationship between myocardial extracellular matrix remodeling and ventricular function, Eur. J. Cardio-Thoracic Surg. 30 (2006) 604–610. doi:10.1016/j.ejcts.2006.07.006.
[99] R.N. Foley, A.M. Murray, S. Li, C. a Herzog, a M. McBean, P.W. Eggers, A.J. Collins, Chronic kidney disease and the risk for cardiovascular disease, renal replacement, and death in the United States Medicare population, 1998 to 1999., J. Am. Soc. Nephrol. 16 (2005) 489–95. doi:10.1681/ASN.2004030203.
[100] A. House, M. Haapio, J. Lassus, R. Bellomo, C. Ronco, Pharmacological management of cardiorenal syndromes., Int. J. Nephrol. 2011 (2011) 630809. doi:10.4061/2011/630809.
[101] F. Genovese, P. Boor, M. Papasotiriou, D.J. Leeming, M.A. Karsdal, J. Floege, Turnover of type III collagen reflects disease severity and is associated with progression and microinflammation in patients with IgA nephropathy., Nephrol. Dial. Transplant. (2015). doi:10.1093/ndt/gfv301.
[102] M. Yamagata, J.R. Sanes, Synaptic localization and function of Sidekick recognition molecules require MAGI scaffolding proteins., J. Neurosci. 30 (2010) 3579–3588. doi:10.1523/JNEUROSCI.6319-09.2010.
[103] X. Wu, K. Hepner, S. Castelino-Prabhu, D. Do, M.B. Kaye, X.J. Yuan, J. Wood, C. Ross, C.L. Sawyers, Y.E. Whang, Evidence for regulation of the PTEN tumor suppressor

by a membrane-localized multi-PDZ domain containing scaffold protein MAGI-2., Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 4233–4238. doi:10.1073/pnas.97.8.4233.
[104] S. Babayeva, Y. Zilber, E. Torban, Planar cell polarity pathway regulates actin rearrangement, cell shape, motility, and nephrin distribution in podocytes., Am. J. Physiol. Renal Physiol. 300 (2011) F549–F560. doi:10.1152/ajprenal.00566.2009.
[105] M.S. Henning, P. Stiedl, D.S. Barry, R. McMahon, S.G. Morham, D. Walsh, M.H. Naghavi, PDZD8 is a novel moesin-interacting cytoskeletal regulatory protein that suppresses infection by herpes simplex virus type 1, Virology. 415 (2011) 114–121. doi:10.1016/j.virol.2011.04.006.
[106] N.S. Eyre, H.E. Drummer, M.R. Beard, The SR-BI partner PDZK1 facilitates hepatitis C virus entry, PLoS Pathog. 6 (2010). doi:10.1371/journal.ppat.1001130.
[107] C. Kranjec, L. Banks, A systematic analysis of human papillomavirus (HPV) E6 PDZ substrates identifies MAGI-1 as a major target of HPV type 16 (HPV-16) and HPV-18 whose loss accompanies disruption of tight junctions., J. Virol. 85 (2011) 1757–64. doi:10.1128/JVI.01756-10.
[108] B.A. Glaunsinger, S.S. Lee, M. Thomas, L. Banks, R. Javier, Interactions of the PDZ-protein MAGI-1 with adenovirus E4-ORF1 and high-risk papillomavirus E6 oncoproteins., Oncogene. 19 (2000) 5270–80. doi:10.1038/sj.onc.1203906.
[109] M. Abrahamson, a J. Barrett, G. Salvesen, a Grubb, Isolation of six cysteine proteinase inhibitors from human urine. Their physicochemical and enzyme kinetic properties and concentrations in biological fluids., J. Biol. Chem. 261 (1986) 11282–11289. doi:10.1002/elps.1150181516.
[110] D. Pende, C. Bottino, R. Castriconi, C. Cantoni, S. Marcenaro, P. Rivera, G.M. Spaggiari, A. Dondero, B. Carnemolla, N. Reymond, M.C. Mingari, M. Lopez, L. Moretta, A. Moretta, PVR (CD155) and Nectin-2 (CD112) as ligands of the human DNAM-1 (CD226) activating receptor: Involvement in tumor cell lysis, Mol. Immunol. 42 (2005) 463–469. doi:10.1016/j.molimm.2004.07.028.
[111] L. Vogt, N. Schmitz, M.O. Kurrer, M. Bauer, H.I. Hinton, S. Behnke, D. Gatto, P. Sebbel, R.R. Beerli, I. Sonderegger, M. Kopf, P. Saudan, M.F. Bachmann, VSIG4, a B7 family-related protein, is a negative regulator of T cell activation, J. Clin. Invest. 116 (2006) 2817–2826. doi:10.1172/JCI25673.
[112] J. Siwy, W. Mullen, I. Golovko, J. Franke, P. Zürbig, Human urinary peptide database for multiple disease biomarker discovery., Proteomics. Clin. Appl. 5 (2011) 367–74. doi:10.1002/prca.201000155.
[113] J. Klein, J. Eales, P. Zürbig, A. Vlahou, H. Mischak, R. Stevens, Proteasix: a tool for automated and large-scale prediction of proteases involved in naturally occurring peptide generation., Proteomics. 13 (2013) 1077–82. doi:10.1002/pmic.201200493.
[114] V. Zavasnik-Bergant, A. Schweiger, T. Bevec, R. Golouh, V. Turk, J. Kos, Inhibitory p41 isoform of invariant chain and its potential target enzymes cathepsins L and H in

distinct populations of macrophages in human lymph nodes., Immunology. 112 (2004) 378–85. doi:10.1111/j.1365-2567.2004.01879.x.
[115] J. Sage, E. Leblanc-Noblesse, C. Nizard, T. Sasaki, S. Schnebert, E. Perrier, R. Kurfurst, D. Brömme, G. Lalmanach, F. Lecaille, Cleavage of Nidogen-1 by Cathepsin S Impairs Its Binding to Basement Membrane Partners, PLoS One. 7 (2012) e43494. doi:10.1371/journal.pone.0043494.
[116] A. Lundwall, V. Band, M. Blaber, J. a Clements, Y. Courty, E.P. Diamandis, H. Fritz, H. Lilja, J. Malm, L.J. Maltais, a Y. Olsson, C. Petraki, A. Scorilas, G. Sotiropoulou, U.-H. Stenman, C. Stephan, M. Talieri, G.M. Yousef, A comprehensive nomenclature for serine proteases with homology to tissue kallikreins., Biol. Chem. 387 (2006) 637–41. doi:10.1515/BC.2006.082.
[117] A. Magklara, A.A. Mellati, G.A. Wasney, S.P. Little, G. Sotiropoulou, G.W. Becker, E.P. Diamandis, Characterization of the enzymatic activity of human kallikrein 6: autoactivation, substrate specificity, and regulation by inhibitors, Biochem. Biophys. Res. Commun. 307 (2003) 948–955. doi:10.1016/S0006-291X(03)01271-3.
[118] A. Wisner, E. Dufour, M. Messaoudi, A. Nejdi, A. Marcel, M.-N. Ungeheuer, C. Rougeot, Human Opiorphin, a natural antinociceptive modulator of opioid-dependent pathways., Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 17979–84. doi:10.1073/pnas.0605865103.
[119] G.I. Rice, D.A. Thomas, P.J. Grant, A.J. Turner, N.M. Hooper, Evaluation of angiotensin-converting enzyme (ACE), its homologue ACE2 and neprilysin in angiotensin peptide metabolism., Biochem. J. 383 (2004) 45–51. doi:10.1042/BJ20040634.
[120] T.G. Yandle, S.O. Brennan, E.A. Espiner, M.G. Nicholls, A.M. Richards, Endopeptidase-24.11 in human plasma degrades atrial natriuretic factor (ANF) to ANF(99–105/106–126), Peptides. 10 (1989) 891–894. doi:10.1016/0196-9781(89)90131-9.
[121] Y. Vanneste, a Michel, R. Dimaline, T. Najdovski, M. Deschodt-Lanckman, Hydrolysis of alpha-human atrial natriuretic peptide in vitro by human kidney membranes and purified endopeptidase-24.11. Evidence for a novel cleavage site., Biochem. J. 254 (1988) 531–537.
[122] N. Morisaki, S. Moriwaki, Y. Sugiyama-Nakagiri, K. Haketa, Y. Takema, G. Imokawa, Neprilysin is identical to skin fibroblast elastase: Its role in skin aging and UV responses, J. Biol. Chem. 285 (2010) 39819–39827. doi:10.1074/jbc.M110.161547.

Figure legends

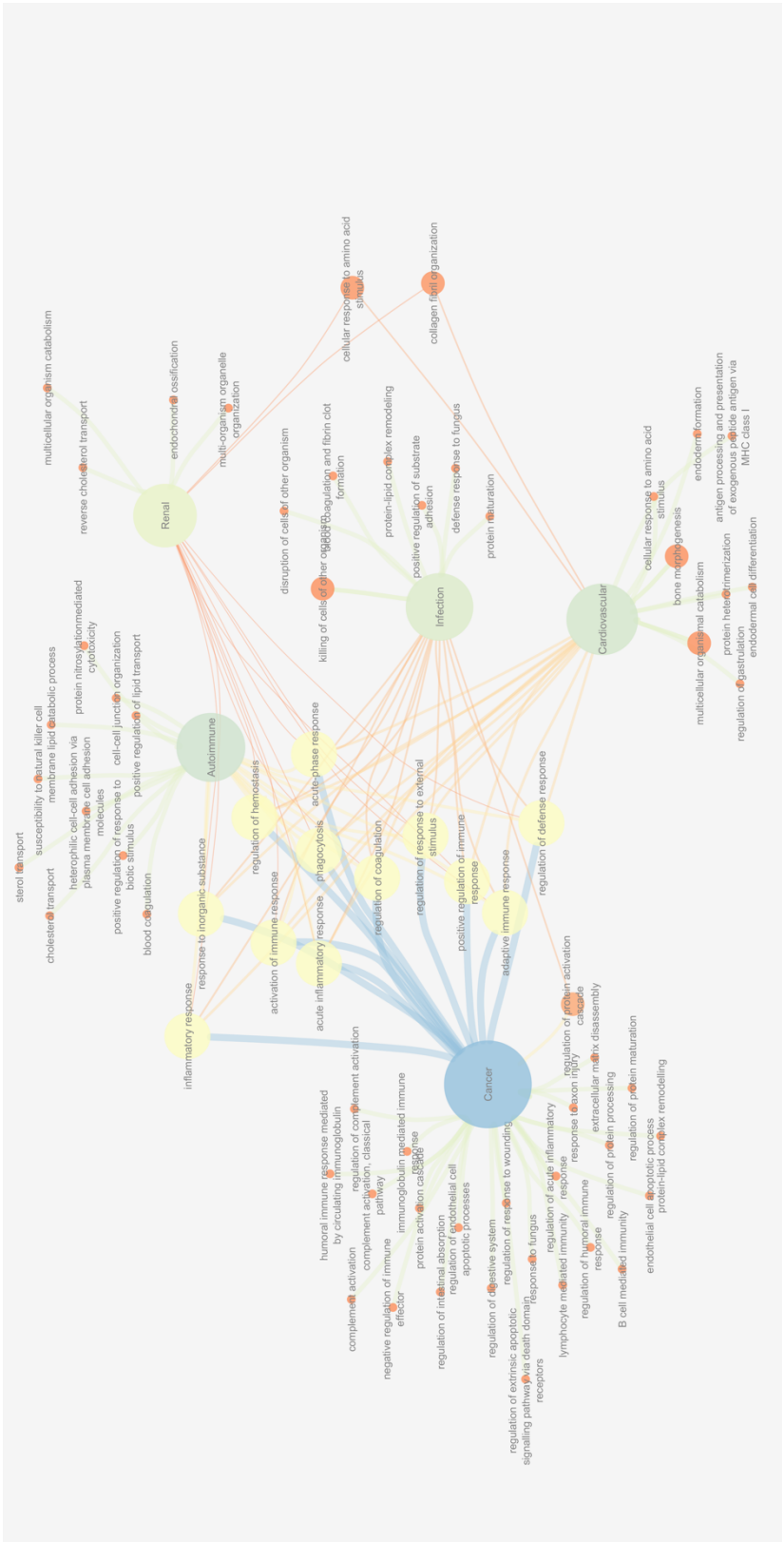


Figure 1: Network depicting betweenness centrality of the main biological

processes represented in the urine of patients across 57 diseases which were grouped by group of diseases. Bigger nodes and larger edges represents its capacity, influence, frequency and connectivity.

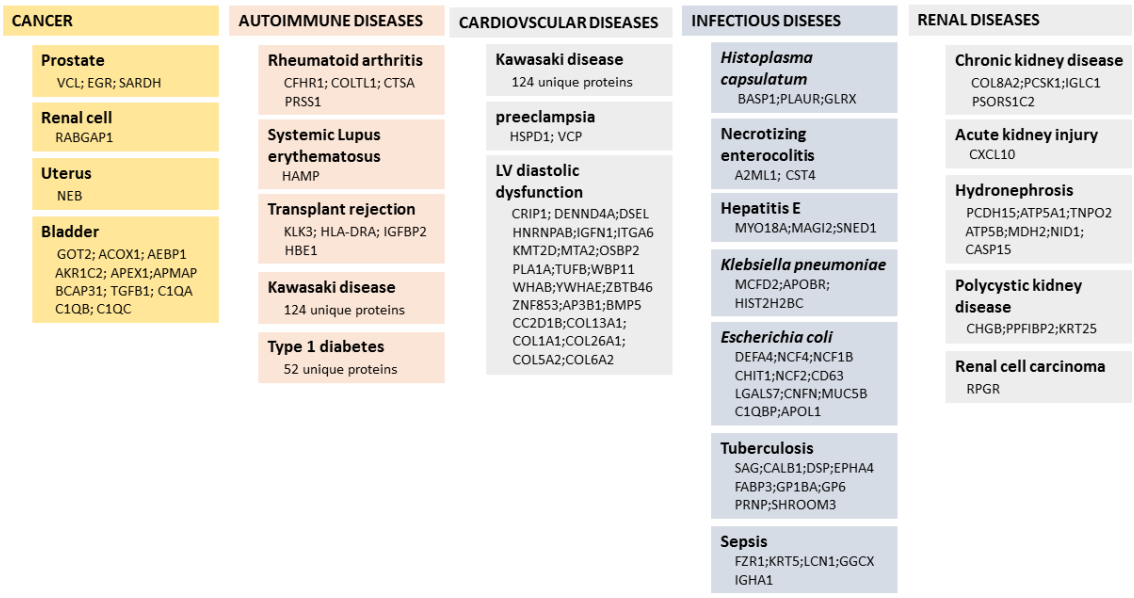


Figure 2: Overview of potential disease biomarkers considering the proteins uniquely overrepresented in the urine.

B – BOOK CHAPTER: *The Role of Urinary Proteases in Bladder Cancer*

Paulo Bastos[1,*] Sandra Magalhães[1,2*] Lúcio Lara Santos[3] Rita Ferreira[2] Rui Vitorino[1,4]

[1]iBiMED – Institute for Research in Biomedicine, Department of Medical Sciences, University of Aveiro, Portugal
[2] QOPNA, Mass Spectrometry Center, Department of Chemistry, University of Aveiro, Portugal
[3] Experimental Pathology and Therapeutics Group - Research Center, Portuguese Oncology Institute - Porto (IPO-Porto), Porto, Portugal.
[4] Department of Physiology and Cardiothoracic Surgery, Faculty of Medicine, University of Porto, Porto, Portugal
*Equal contributors

Corresponding author:

Rui Vitorino
Department of Medical Sciences
University of Aveiro

e-mail: rvitorino@ua.pt

NOTE: This book chapter has been accepted to integrate the book entitled "**Pathophysiological Aspects of Proteases**" (chapter 4) upon invitation by the **editor Springer**. It is currently *In Press*. ID number: 26256699

Abstract

Bladder cancer (BCa) is one of the most prevalent malignancies worldwide. Risk factors for BCa are well established and include smoking and infections, which can lead to immune system activation, altered gene expression patterns, proteolytic activity, tissue damage and, ultimately, to cancer development. Urine has become one of the most attractive diagnosis samples and, notably, urine profiling by mass spectrometry allows the simultaneously analysis of multiple enzymes and their interactors, substrates, inhibitors and regulators, providing an integrative view of enzymatic dynamics. Most BCa-associated enzymatic alterations take place at the level of proteases, being MMP-9, MMP-2, urokinase-type plasminogen activator, cathepsin D and cathepsin G already related to BCa development and progression. Herein, we overview the role of proteases and the classes more studied in BCa pathogenesis, as well as the methodologies used for assess proteases amount and activity in urine samples, highlighting its advantages and limitations, and the value of urinary proteases as disease biomarkers.

Introduction

Bladder cancer (BCa) is the ninth most common cancer worldwide with an estimated incidence above 400,000 new cases *per* year. BCa incidence rises with aging and is more frequent in men, which account to be the seventh most common cancer in this gender [1]. In women, BCa is the seventeenth most common type of malignant disease [1]. Whereas men are three to four times more likely to develop BCa, women have worse prognosis [2, 3]. This sexual dimorphism may be related to differences in the exposure to carcinogens such as tobacco or to genetic, hormonal and anatomical factors [2]. There are several risk factors involved in the pathogenesis of this disease [4]. According to World Health Organization (WHO), cigarette smoking and exposure to aromatic amines are the most important risk factors [5, 6]. In fact, the risk of developing BCa in smokers is about 2 to 6 times higher compared to non-smokers [5]. In turn, occupational exposure to carcinogens like aromatic amines is responsible for 25% of all BCa cases [5–7].

BCa is more frequent in developed countries where, in 2012, approximately 10 among every 100,000 individuals were diagnosed with BCa. Among developed regions, higher incidence rates have been reported in both Europe and North America [1]. As a consequence of its high incidence, prevalence and aggressiveness, BCa is one of the most expensive diseases to treat and to manage [8–11]. Most of this economic burden results from the high number of invasive procedures performed (e.g. cystoscopies) and from the several hospitalizations required for its management [8–11]. Taken together, incidence, mortality, prevalence and costs of BCa make this disease a public health problem, even though it is not officially regarded as such.

Currently, BCa is classified as a multifactorial disease that can be divided into three histological subtypes: urothelial cell carcinomas, adenocarcinomas and squamous cell carcinomas [5, 12]. In developed countries, namely United States, France and Italy, urothelial cell carcinomas constitute about 90% of BCa cases, while the incidence of the remaining types of BCa is much lower, ranging from 1.1 to 2.8% for squamous cell carcinomas and from 1.5 to 1.9% in the case of adenocarcinomas [5]. Urothelial cell carcinomas can be classified into low-grade BCa and high-grade BCa, a classification that is based on the histological analysis of the bladder. Low-grade BCa is less aggressive and

rarely invades bladder muscle tissue or metastasizes to other parts of the body, and patients rarely die from it. On the other hand, high-grade BCa commonly invades muscle wall and metastasizes, therefore requiring more aggressive treatments. Consequently, high-grade BCa is responsible for almost all BCa-related deaths [7]. Another way to classify BCa is based on the invasion of the bladder muscular layer. Nonmuscle-invasive BCa is often treated by the surgical removing the tumors, sometimes combined with localized chemotherapy [7, 13]. Patients with nonmuscle-invasive BCa have a 5-year survival rate ranging from 82 to 100% for early stage BCa. On the other hand, muscle-invasive BCa is much more aggressive, and treatment involves surgical removal of the bladder as well as aggressive radiotherapy and chemotherapy treatments [7, 13]. Even with the complications associated with surgery, surgical intervention for tumor removal remains the most effective treatment, highlighting the lack of specific therapeutic targets and effective pharmacological agents. However, even complete resections are followed by recurrence rates of 70% for superficial bladder tumors and intravesical chemotherapy can attenuate this extremely high recurrence rate by only 15% at most. Nevertheless, this type of BCa has a poorer prognosis, so that only 5 to 10% of patients with metastasis live more than 2 years after diagnosis [14].

The proportion of squamous cell and transitional cell bladder carcinomas has been changing, at least in some populations, most likely due to increased exposure to distinct etiologic factors, especially smoking [15]. In areas where schistosomiasis is endemic (e.g. Sudan, Egypt), squamous cell carcinoma account for 75% of all malignant bladder tumors. However, the increased smoking prevalence has been responsible for a significant shift towards urothelial/transitional cell carcinoma (6-fold increase over squamous cell carcinoma) [15, 16]. This observation has important repercussions as urothelial/transitional cell carcinoma patients are poorer candidates for surgical intervention and present increased risk of recurrence [16]. In addition, other risk factors for developing BCa include family history of the disease, genetic mutations in *HRAS*, *Rb1*, *PTEN/MMAC1*, *NAT2* and *GSTM1* genes, exposure to arsenic, and certain drugs used in chemotherapy like cyclophosphamide [5, 7].

There are several pathways involved in bladder tumorigenesis. At a molecular level, it is known that mutations in *FGFR3, TP53, RB1, ERBB2 and PTEN* genes are involved in

bladder carcinogenesis [17, 18]. Also, some oncogenic miRNAs are upregulated and tumor suppressor miRNAs are downregulated in BCa patients [17]. At the metabolome level, like in other types of cancer, the Warburg effect contributes to BCa development, progression and aggressiveness [19]. Neoplastic cells have a strong dependence on glycolysis in conditions of normoxia, and BCa cells display an upregulation of genes involved in glucose uptake, like GLUT1 and GLUT3 transporters and Akt, which promotes the PI3K/Akt/mTOR pathway [19]. There is also an overexpression of pyruvate-related and pentose phosphate pathway-related genes, and glycogen metabolism is enhanced during BCa development. Simultaneously, there are alterations in lipid metabolism, like increased fatty acid synthesis and oxidation, changes in the tricarboxylic acid cycle and ketogenesis activity [19], as well as the hyperactivation of proteases, which may lead to tissue damage and invasion, immune system and apoptosis evasion, tumorigenesis and cancer metastasis [20, 21].

Early detection of BCa is not always easy, since BCa is commonly asymptomatic at the early stages. Also, the first sign of BCa is usually painless hematuria, which is shared by other diseases [5, 12, 18]. Moreover, other symptoms may include abdominal pain, fatigue, weight loss and urinary frequency and urgency. Thus, diagnosis of BCa can be performed by ultrasounds, computed tomography scans and non-invasive urine cytology [5, 12]. Urinary cytology has the advantage of being a non-invasive approach, but despite its high specificity, it lacks sensitivity to detect low-grade tumors, so that cystoscopy is always done to corroborate the diagnosis [5, 12].

One of the biggest and most promising challenges in BCa research is the development of a diagnostic tool based on biofluids' profiling. Not surprisingly, urine is the most promising biological fluid, once it can be collected in high amounts by a non-invasive and simply procedure [12]. Also, in the case of BCa, urine better reflects alterations taking place in the bladder because it is in direct contact with the tissue [12]. Several studies have explored the diagnosis value of biomarkers for urine assay in BCa [22–25]. Among the identified putative biomarkers are a few proteases, reflecting the role of proteolytic systems in carcinogenesis and metastization [22, 23]. However, the huge interplay between multiple proteases and its inhibitors makes the assay of proteases in urine complex and unpredictable. In this chapter, we will explore the role of proteases

in BCa pathophysiology and how these enzymes might be assessed in biofluids for the diagnosis and management of this disease.

An Update of the existing biomarkers for the clinical management of BCa

Currently, there are three FDA-approved urinary tests for the screening and surveillance of BCa. NMP22 BladderCheck Test® is an enzyme immunoassay for the detection of NMP22, a nuclear matrix protein recognized as an urothelial-specific cancer biomarker [26, 27]. This test has the advantage of being inexpensive, rapid and operator-independent. However, it leads to a high number of false positives [23, 27]. The FDP test is another FDA-approved one that detects increased urinary concentration of fibrin/fibrinogen degradation products (FDPs), which are associated with malignant tumors [23, 27]. However, similarly to the NMP22 test, lacks sensitivity (68%) and displays a high rate of false positives, particularly in patients with hematuria. So, it was recently removed from the market [23]. Lastly, BTAstat and BTA-TRAK are two tests that detect the bladder tumor antigen (BTA) in urine. BTA is the human complement factor H-related protein and it is produced by human BCa cells but not by normal epithelial cells [26]. These tests were approved for surveillance only, since both display a high rate of false positives. Despite its high sensitivity for low-grade lesions, the sensitivity of BTA-tests for high-grade lesions is worse than cystoscopy [23]. Even so, it is possible to improve the sensitivity of these tests by combining them, but the rate of false positives remains high [26].

Therefore, efforts continuous to be made for the identification of biomarkers that help to improve the clinical management of BCa. Still, there is a general awareness that it will be difficult, if not impossible, to identify a single biomarker with high specificity and sensitivity (a golden bullet). Consequently, emphasis has been given to the combination of multiple protein markers that when analyzed together in a multimarker panel may allow the diagnosis of BCa and improve the clinical management of this type of cancer. Considering the role of the proteolytic systems in bladder carcinogenesis and metastization, proteases might be seen as markers to be included in such diagnosis panels. Because proteolytic activity is regulated at multiple levels, including gene expression, enzyme compartmentalization, regulation by modulators and degradation

[28], the biological significance of urinary proteases and quantitative alterations are difficult to interpret and not readily predictable. The urinary levels of primarily intracellularly-acting proteases may not accurately reflect local mRNA expression since these proteases may be enriched in the tissues and poorly secreted in urine [29–40]. In turn, urinary levels of highly secreted proteases are subjected to the influence of multiple factors including the rate of secretion, tumor burden and transcription regulation [41–43]. As such, the exploitation of quantitative changes in the urinary protease levels has revealed a far more challenging avenue than initially envisioned. Still, proteomics appears to be the most promising approach to characterize urinary proteases and promises to allow BCa screening, diagnosis, monitoring, therapy management and the identification of novel pharmacological targets.

Urine Proteases profiling for biomarker screening

When studying proteins and peptides in urine samples, different enrichment procedures are usually applied, given different but complementary results. The identification and characterization of putative urinary biomarkers may be performed with gel-based or gel-free approaches (Table 1). Gel-based approaches usually involve protein separation by 1D/2D Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) whereas gel-free approaches includes Capillary Electrophoresis (CE) or Liquid Chromatography (LC). Nevertheless, gel-free and gel-based approaches might be combined with each other into a strategy known as GeLC. Starting from a complex sample such as urine, 1D/2D SDS-PAGE allows accurate separation of proteins according to their molecular weight and/or isoelectric point, resolving hundreds of proteins, then subjected to MS-based analysis, immunoblot detection and/or zymography-based activity assays (Table 1) [29–31, 34, 36–38]. One of the main shortcomings of gel-based approaches results from the high-salt concentrations in urine (mainly urea), as well as the presence of lipids and glycosaminoglycans, which interfere with proteomic analysis (particularly during isoelectric focusing in 2D SDS-PAGE). Despite the versatility of gel-based approaches, these may become daunting when dealing with complex samples, particularly for high-throughput analyses. Consequently, LC has become the platform of choice for protein separation from urine samples of patients with BCa, providing better resolution and

allowing easier sample manipulation (Table 1) [30, 32, 33, 35–37]. Furthermore, distinct chromatography-based techniques such as dual-lectin affinity, strong-cation exchange and reverse phase chromatography exploit unique physicochemical properties, which allow the enrichment of particular urinary sub-proteomes and the identification of novel enzymes as well as enzyme inhibitors [24, 32]. Similarly to gel-based strategies, gel-free approaches can also be followed by MS analysis (with or without prior digestion), immunoblotting and/or zymography (Table 1).

Immunoaffinity assays

Immunoaffinity or antibody-based approaches can be used either to isolate intended targets or to deplete unwanted ones, and these techniques have been exploited during all steps of urinary biomarker discovery, from targeted sample depletion and protein isolation (e.g. immunoprecipitation) to biomarker validation (e.g. western blot, Dot blot) [30, 33, 44]. In addition to be highly sensitive and specific, immunoaffinity assays are, above all, the methods of choice required for biomarker validation. Particularly, as will be discussed in the subsequent sections, immunoaffinity assays such as western blot have been successfully used for biomarker validation using urine samples from BCa patients, including the proteases cathepsin D, cathepsin G, cathepsin L, matrix metalloproteinase-2 (MMP-2), matrix metalloproteinase-9 (MMP-9), A disintegrin and metalloproteinase with thrombospondin motifs 7 (ADAMTS7), metalloproteinase inhibitors (e.g. TIMP1), among others.

In recent years, classical blotting techniques have started to be replaced by surface plasmon resonance imaging, which employ antibody arrays for the multiplexed detection of protein biomarkers [45, 46]. This technology have been applied to the detection of BCa in its early stages, through the assessment of cathepsin D/protein ratio in both serum and urine [47].

Zymography

Zymography has assumed a leading role when studying urine proteases because it allows protein resolution based on molecular weight and isoelectric point, enzymatic activity analysis and the screening of substrates and enzyme inhibitors. Despite this

versatility, zymograhy-based studies of proteolytic alterations in BCa patients have relied almost exclusively on PAGE-based resolution of urinary proteins, incubation with a substrate more or less specific for a group of proteases (e.g. gelatin zymography to access the activity of MMP-2 and MMP-9) and estimation of enzymatic activity by substrate hydrolysis quantification [48–50]. Also, zymography allows multiple analysis to be performed, either simultaneously or sequential, thus largely expanding the repertoire of possibilities for analysis [51, 52].

By analyzing enzymatic activity levels, when applied to specimens collected from BCa patients, zymography has provided biomarker-based models with specificities of up to 100%, identifying molecular players not amenable to be monitored by other techniques [48–50, 53]. The concentration of urinary proteases may not accurately reflect proteolytic activity [48, 49]. Notably, zymography allows to study and to quantify enzymatic activities in addition to enzymatic concentrations, thus providing information regarding quantitative changes as well as qualitative alterations. As such, zymography becomes and invaluable tool for the analysis of functional aspects, as seen for BCa where enzymatic trafficking is impaired and protease/inhibitor complexes or protease dimers may be frequent, with repercussions to the proteolytic function [48–50, 53].

Peptidomics

The screen of native peptides (peptidomics) in urine, particularly polar, charged or chargeable small ones, has mostly relied on CE analysis. CE-based urine peptidomics aims to screen the peptides differentially expressed across distinct populations, providing an indirect window to assess the proteases modulated by BCa. Such indirect urine profiling approach has been proven capable of accurately discriminate cancer patients from patients with other organ-related nonmalignant conditions [54]. Together with bioinformatics tools, CE may serve for the discovery and validation of multimarker panels [55–57], and for the comparison of excreted urinary peptides to endogenous “housekeeping” peptides [55, 56]. CE-based native urinary peptides profiling has been applied for the detection of primary (with up to 91% sensitivity and 68% specificity) and recurrent (with up to 87% sensitivity and 51% specificity) urothelial BCa [57].

Peptides shown to most accurately allow the detection of BCa were collagen fragments (~60%), distantly followed by hemoglobin, apolipoprotein A-I peptides, membrane-specific heparin, sulfate proteoglycan core protein, A disintegrin and metalloproteinase domain-containing protein 22 (ADAM22) and A disintegrin and metalloproteinase with thrombospondin motifs 1 (ADAMTS1). Together, these peptides reflect increased proteolytic activity, extracellular matrix remodeling, collagen degradation and hematuria [57].

However, there are some sampling issues that should be considering in the analysis of proteome data retrieved from proteome profiling of urine samples collected from patients with BCa. Urine proteomics applied to the search of BCa biomarkers have focused most extensively on older patients (over the age of 60) [32, 34, 36], which reflect the clinical phenotype of the majority of patients but reduces the likelihood of finding markers for early diagnosis purposes. Gender also influences the urine proteome profile. Females tend to present higher albumin and transferrin urinary concentration and a more complex urinary subproteome (composed by the less abundant proteins) [58]. Therefore, the pathophysiological role and the prognosis and diagnosis value of a given BCa-related urinary biomarker may not be suitable across all age groups or across genders.

URINARY PROTEASES ASSOCIATED TO BLADDER CANCER

Urine has become one of the most attractive diagnosis sample, involving noninvasive collection of unrestricted quantities [59]. Generally, less than 150 mg of proteins is excreted in urine, mainly due to glomerular filtration, and partially by tubular reabsorption, secretion, and degradation [60, 61]. Around 70% of urinary proteins are of kidney origin [62], a percentage modulated by diseases. In addition to kidney diseases, recent findings show that each and several diseases can be accurately discriminated by a unique protein fingerprint. From the more than 2,500 proteins identified by MS-based urinalysis as differentially expressed across human diseases, approximately 15% (approximately 360 proteins) result from malignant diseases, particular BCa (58% or 209 proteins of these approximately 360 malignancy-associated differentially expressed urinary proteins) [63, 64]. Indeed, by pooling data retrieved

from ten MS-based studies concerning differentially expressed urinary proteins from patients with BCa [29–38], 209 proteins were highlighted. Among these, 44 proteins were classified as enzymes, 21 of which are hydrolases and from these 13 are proteases (Tables 2 and 3). These proteases/peptidases are listed in Table 3 together with its function and expression variation (up/down) in BCa. Likewise, the expression patterns of several genes have been determined by microarrays in close to 1,500 BCa patients and revealed several proteases as putative diagnostic and prognostic BCa markers (Table 3) [65–75]. However, as also depicted by Table 3, alterations in proteases mRNA do not always reflect those concerning the concentrations of proteases neither their activity. For instance, while both levels and the activity of MMP-2 and cathepsin G are upregulated in urine samples from BCa patients, their mRNA expression levels are downregulated (Table 3), which may reflect increased stimulation-dependent activation (by, for instance, inflammatory and stress-response signaling pathways), stimulation-independent activation or enhanced intrinsic hydrolytic activity or alterations in enzymatic/lysosomal trafficking (see the following section and Table 3).

The role of proteases goes behind protein digestion with recognized important functions in several signaling pathways (Tables 2 and 3, Figure 1 and 2), namely those involved in DNA replication and translation processes, cellular proliferation, tissue remodeling, angiogenesis, fertilization, hemostasis, blood coagulation, apoptosis, necrosis, immunity, organelle recycling and cellular senescence [76–78]. Nevertheless, the human “degradome”, which results from protein digestion, has been suggested to play a critical role in the pathogenesis of some diseases. Even thought, only 2% of protein coding genes contributes to such “degradome” [76–84]. These proteases are divided into cysteine proteases, serine proteases, aspartic proteases, threonine proteases or metalloproteases [85]. Since the scope of this chapter is the biological role and clinical implications of urinary proteases in BCa, a brief description of proteases’ families follow.

Cathepsins

Cathepsin proteases comprises a diverse family of hydrolytic enzymes, which are expressed at the cell surface and/or secreted into the extracellular milieu. These

proteases are classically known for its role in lysosomal protein turnover and extracellular matrix degradation [86]. Despite this classical view, cathepsins encompass many and much more diverse functions. For instance, cathepsin S plays an important role in MHC-II-mediated antigen presentation, so that animal models lacking this enzyme cannot activate the MHC-II, displaying markedly impaired antigenic peptide binding, antibody class switching and splenocytes/dendritic cells function [87]. Moreover, while trafficking through the endocytic pathway to allow invariant chain degradation and antigenic peptide loading onto MHC-II may be mediated by either one of several cathepsins in mice, this redundancy in humans is not observed and, therefore, alterations in the levels of one particular cathepsin may have more severe consequences than in animal models [88]. The involvement of cathepsins on progressive tissue remodeling [89] is also required for malignant diseases development, progression and metastasis [86, 90], which might explain the upregulation of these enzymes in BCa (Table 3).

Most members of this family of proteases (B, C, F, H, L, K, O, S, V, W and X) display proteolytic activity at cysteine residues, while cathepsins A and G are serine carboxypeptidases and cathepsin D and E are aspartic proteases [91]. In turn, the activity of these proteases is regulated by its corresponding inhibitors cystatins and serpins [91]. Upregulated in the urine from BCa patients (Figure 1), cathepsin G is a serine peptidase highly expressed in immune cells, most notably in neutrophils and mast cells, but also in monocytes and dendritic cells [92, 93]. Therefore, the high levels of cathepsin G, detected by MS profiling of urine from BCa patients (Table 3), is most likely of immune cell origin (rather than tumor cell origin) and may account for some of the biological alterations observed in BCa patients, such as NF-kappa B signaling pathway, eicosanoid synthesis, leucocyte transendothelial migration and transcriptional deregulation (Figure 2, Table 3).

Cathepsin G is known to cause extensive damage to host tissues due to its poor specificity [94]. Animal models of ischemia lacking cathepsin G can survive, while the condition is lethal to most wild type animals. For instance, while neutrophil infiltration and the levels of CXCL1 and CXCL2 chemokines and myeloperoxidase are equal in both groups immediately after induction of ischemia, mice lacking cathepsin G do not suffer

from the severe tubular necrosis, tubular cell apoptosis and fibrosis characteristic of normal mice, showing that cathepsin G sustains tissue pathology and fibrosis after stress, inflammation and injury induction [94].

Cathepsin G derived from neutrophils is co-released into phagolysosomes, thereby helping to fight of infections [95]. In fact, studies with animal models suggest that cathepsin G activation is required for survival upon infection by *Aspergillus fumigatus* and *Staphylococcus aureus* [95, 96]. Likewise, both membrane-bound and secreted forms of cathepsin G co-localize and are co-released with other peptidases [97, 98], possessing antimicrobial activity and being the main neutrophil-derived antibacterial mediators [95]. Accordingly, its antimicrobial activity seems to be partially enhanced by and dependent on other peptidases like elastase [95], and its surface expression may increase up to 30-fold upon stimulation [97]. Therefore, infections may trigger the upregulation of cathepsin G as a host protective mechanism, but may also cause tissue damage and promote tumorigenesis (Figure 2) or metastization once cancer has been established.

Compared to other human proteases, cathepsin G is less selective by combining both chymotryptic and tryptic hydrolytic activities [99], and while mutations/evolutionary alterations responsible for this duality confer it a broader activity spectrum, these may also contribute for its lower potency [100]. Irrespectively, its broad spectrum of activity most likely contributes for extracellular matrix degradation, complement activation, proteoglycans degradation and lysosomal signaling alterations in BCa (Figure 2, Table 3). It should be noted that most of what is known about cathepsin G activity results from *in vitro* assays or animal studies, which may hamper the interpretation of its hypothetical implications in BCa. For instance, in contrast with human cathepsin G, murine cathepsin G displays a chymotryptic bias and lacks tryptic activity to a large extent, is more specific, but significantly more potent [100].

The major activator of cathepsin G and related peptidases is the lysosomal cathepsin C (CTSC) or dipeptidyl peptidase I (DPEP1) [101], which was also found in increased levels in the urine from BCa patients (Figure 1, Table 3). Increased levels of DPEP1 and the corresponding activated peptidases are associated with exaggerated immune responses and tissue damage. In contrast, animal models lacking DPEP1 are protected from these

effects upon the induction of inflammatory response and display decreased expression of TNF-α and IL-1β [101], also involved in MMP-mediated BCa progression as discussed in the following section. Its tryptic activity, in turn, activates proteinase-activated receptors, calcium mobilization and neutrophil-platelet interactions at sites of injury or inflammation [102], as well as complement proteins [103] and pro-urokinase plasminogen activator (PLAU) [104], which are in high levels in the urine of BCa patients (Table 3). Owing to its peptidase activity, human cathepsin G hydrolyzes angiotensin-(1-10) (angiotensin I) to yield angiotensin-(1-8) (angiotensin II) [105, 106] and activates metallopeptidases [107], thereby accounting for the alterations observed in the renin-angiotensin and lysosomal signaling (Figure 2, Table 3). Such metallopeptidases include MMP-2, which is also in high content in the urine from BCa patients (as described in the next section) and the regulation of both substrates may therefore be intimately interconnected [107]. Taken together, cathepsin G, as well as its promoter and its substrates are increased in urine samples from BCa patients, and seem to be involved in the promotion of extracellular matrix components degradation (e.g. collagen, laminins) and local tissue damage, allowing malignant cells to infiltrate the bladder tissues and to disseminate into the blood circulation.

In contrast with cathepsin G (a serine protease), cathepsin D is an aspartyl protease involved in cellular components' recycling and apoptosis control. Procathepsin D is found in the Golgi complex and is proteolitically inactive but the intermediate and mature forms of cathepsin D are enzymatically active [108]. These forms are found in endosomes and lysosomes, respectively, and are responsible for the autophagy and apoptosis pathways [109], thus accounting for the enriched processes illustrated in Table 3 and Figure 2. For instance, cathepsin D is known to function as an anti-apoptotic mediator in human malignant glioblastoma cells, inducing autophagy under stress conditions and conferring cancer cells resistance against genotoxicity mediated by oxidative agents such as hydrogen peroxide [110]. However, the proteolytic activity of cathepsin D also allows the release of growth factors that act by promoting tumor cell growth [109]. Cathepsin D mediates both metastasis and recurrence in breast cancer [111, 112] and metastasis/dissemination in laryngeal and pancreatic cancers [113, 114]. It was proposed as a marker of vascular and microvascular density in breast and ovarian

tumors [115, 116]. However, in contrast with cathepsin G, which appears to be mostly of immune cells' origin, cathepsin D has been described as secreted primarily by cancer cells [117, 118]. In a small cohort of patients with BCa (15 patients), it was shown that the activity of cathepsin D was significantly increased in serum samples [119]. The urinary concentration of cathepsin D is approximately equal to that of serum in both BCa patients and healthy subjects [47]. In another small cohort of patients (21 subjects), tissue cathepsin D analyzed by immunohistochemistry displayed a significant but inverse correlation with tumor grade and stage. This data reinforce the involvement of this protease in the early stages of the disease and local tissue invasion, but its expression decrease once high-grade and high invasiveness are attained [120]. A disengagement between the regulation of cathepsin D and its inhibitor cystatin C was also reported [120]. Subsequently, in a more comprehensive study comprising 68 BCa patients, both serum and urinary concentrations of cathepsin D were found significantly higher, using the Surface Plasmon Resonance Imaging (SPRI) biosensor [47]. The serum cathepsin D/creatinine ratio was reported to be significantly higher in superficial tumors (Ta + T1) compared to invasive tumors (T2 + T3). This observation makes cathepsin D of particular value for detecting BCa in its early stages. Of uttermost importance, urinary cathepsin D/protein ratio was significantly higher in primary *vs* recurrent (2.24-fold) and low-grade *vs* high-grade (1.67-fold) tumors, reinforcing its diagnosis value for BCa in the initial stages [47]. Overall, BCa displayed significantly higher serum (8-fold) and urine (7-fold) cathepsin D concentrations when compared with healthy controls, and differences remained significant even when normalized to total protein and creatinine levels [47].

Similarly to cathepsins D and G, urinary cathepsin L is present in significantly higher urinary levels in patients with urothelial carcinoma compared to normal controls, even in patients with negative cytology [40]. Urinary cathepsin L levels were higher in patients with higher grade, being significantly associated with tumor invasiveness. Also, the associations with BCa presence and aggressiveness were maintained after adjusting cathepsin L levels for urinary creatinine content [40]. Perhaps of uttermost value, voided urinary cathepsin L was proven to be an independent predictor of BCa recurrence and invasiveness at advanced stages, outperforming both cytology and NMP22 (an FDA approved marker for BCa at initial stages) [40]. These results were in accordance with

increased cathepsin L mRNA expression in tissue samples from BCa patients [39]. Cathepsin L expression in tumor cells is under distinct control mechanisms compared to normal cells, as the transcription factors responsible for basal cathepsin L expression (NF-Y, Sp1, Sp2 and Sp3) are not responsible for its overexpression in tumor cells [121]. The 5' region of the cathepsin L gene encompasses CpG islands, which regulate promoter activity, and demethylation of these CpG islands seems to be positively associated with cathepsin L-dependent types of malignant diseases, but not independent ones [121]. Therefore, one possible mechanism (albeit not tested to date) contributing for the upregulation of cathepsin L in BCa may encompass epigenetic alterations, particular DNA demethylation, explaining the transcriptional misregulation depicted in Table 3 and Figure 2.

Overall, the upregulation of cathepsins and their inhibitors is known to mediate multiple stages of tumorigenesis, carcinogenesis and angiogenesis [90, 122]. In addition to directly carrying out extracellular matrix and basement membrane degradation, cathepsins also disrupt intercellular adhesion proteins (e.g. E-cadherin, at adherens junctions) and activate proteolytic cascades in which the activity of other proteases such as MMPs and urokinase plasminogen activator are enhanced [86, 122]. Therefore, the levels of cathepsins and, more specifically, the cathepsin/CIP (cathepsin inhibitor) ratio, may be used as a suitable marker of malignancy [39]. Nevertheless, among the cathepsin protein family, cathepsins G and D seem to be the best potential urinary biomarkers for BCa (Figure 1, Table 3).

Matrix Metalloproteinases

MMPs are zinc-dependent endopeptidases crucial for many cellular processes, from physiological and developmental to pathological ones [123, 124]. Accordingly, these enzymes are particular important in regulating the local cellular and tissue microenvironment, not only by carrying out extracellular matrix remodeling but also by mediating several physiological processes through signaling regulation [125]. There are 23 human MMPs catalogued to date, most of which display the conserved pro-peptide, catalytic, and hemopexin-like C-terminal domains. Accordingly, MMPs are expressed as inactive pro-preptides with their zinc ions in the catalytic sites attached to the pro-

peptide domain, being activated by the cleavage of this domain. Even though the cleavage of the pro-domain requires the action of another proteolytic enzyme (intracellular furin or extracellular MMPs/serine proteinases), modifications taking place at the cysteine residue of the pro-peptide domain may also compromise such interactions and thus activate MMPs [126, 127].

In the setting of BCa, uncontrolled proliferation and compromised apoptosis is attributed in a large extent to the upregulation of zinc-dependent endopeptidades MMP-2 and MMP-9 (Figures 1 and 3, Table 3) [128]. The proteolytic activity of MMPs is responsible for the cleavage of ligands and corresponding receptors responsible for conveying proapoptotic signals, thus allowing tumor cells to evade apoptosis [129]. These changes mediate both tumor cell migration and the development of new tumor-related blood vessels [53, 130]. So, MMP-2 as well as MMP-9 have been regarded as promising biomarkers for BCa [21, 49, 131, 132]. MMP-9 acts synergistically with ADAM metallopeptidases and epidermal growth factor receptor (EGFR) in order to degrade E-cadherin and liberate β-catenin, allowing its translocation into the nucleus and thus enhancing cellular proliferation [133, 134]. Also, tumor cells may use transforming growth factor-β (TGF-β) signaling via MMP-mediated proteolytic conversion, allowing tumors to evade immune surveillance and thus leading to local invasion and cancer metastasis [135, 136]. Similarly, ADAM-mediated proteolytic cleavage of tumor-associated major histocompatibility complex class I-related proteins MICA and MICB, which suppresses NK cell-mediated cytotoxicity, seems to require the activity of MMPs [137, 138]. Inflammatory cell-derived MMP-9 increases the bioavailability of vascular endothelial growth factor (VEGF), which is the most potent inducer of tumor angiogenesis [139].

High molecular weight forms of MMPs are also found in increased levels in the urine of BCa patients, most notably MMP-9 dimers and MMP-9 conjugated with its inhibitor TIMP1 [53], and MMP-2 conjugated with the MMP inhibitor TIMP2 [132]. The overexpression of MMP-9 promotes the upregulation of its inhibitor TIMP1 as a countermeasure. However, when MMP9 is in excess over its inhibitor TIMP1, the formation of MMP-9 disulfide-bonded dimers takes place, forming a more stable and slowly activating MMP-9, which is less susceptible to inactivation [140]. MMP-9/NGAL

conjugates seem to protect MMP-9 from autodegradation and are associated with the formation of cancer metastasis [53, 141, 142]. MMP-2, MMP-9, MMP-9/NGAL complex and MMP9- dimer are overexpressed in high grade BCa patients compared to low grade ones, and MMP-9/TIMP1 complex is overexpressed in larger tumors compared to smaller ones [48]. Also, MMP-2 monomer, MMP-2 conjugated with its inhibitor and MMP-2 fragments are significantly associated with high grade/*lamina propria* invasion [132], while the presence of MMP-9/TIMP2 discriminates malignant tumors from benign ones [50]. The misbalance given by the ratio MMP/MMP inhibitor, rather than upregulation of any MMP independently, seems to be characteristic of BCa [30]. Moreover, in addition to enzyme monomers and enzymes complexes, immunoblotting approaches have revealed the presence of multiple MMP-2 fragments in the urine of BCa patients, which are also associated with the disease, particularly with transitional cell carcinoma cases [132].

More than measuring the levels of MMPs, MMPs' proteolytic activity assessed by zymography has assumed a paramount role in discerning the involvement of MMPs in BCa. In these studies of enzymatic activity, close to 65% of BCa patients present increased urinary proteolytic activity compared to healthy controls [48, 49]. The assessment of the proteolytic activity of multiple MMPs as provided specificities of up to 100% for BCa detection [48, 50]. Most of the proteolytic activity characteristic of urine samples from BCa patients seems to be carried out by MMP-9, followed by MMP-9/NGAL conjugates, MMP-2 and MMP-9 dimers. When MMP-9 dominates, the maximum sensitivity resulting from urine MMPs profiling for BCa detection seems to be attained with MMP-9 (62.1%), followed by MMP-9/NGAL conjugates (60.9%), MMP-2 (54.5%) and MMP-9 dimers (53%) [79]. Irrespectively, MMPs are positively correlated with each other. So, the increased levels of a particular MMP upregulates the remaining one. In this regard, MMP-9 seems to suffer the highest positive influence, being more sensitivity and more susceptible to a larger number of factors [79,141]. In fact, the expression of MMP-9 is enhanced by inflammatory cytokines such as TNFα and interleukins, being its increase more susceptible to alterations triggered by inflammatory conditions (e.g. cystitis) and less specific to BCa itself [156], as illustrated by patients with cystitis who excrete significantly more MMP-9 but not more MMP-2

than controls [141]. In contrast, MMP-2 allows a better discrimination between BCa patients and healthy subjects [139,156], being more specific and less influenced by cancer-unrelated factors [79]. However, MMP-2 and MMP-9 have failed to accurately identify BCa in the early stages/superficial bladder carcinoma (stages CIS, Ta and T1). Urinary MMP-2 seems so be a suitable biomarker for high BCa tumor grade (G3), while both MMP-2 and MMP-9 are suitable for advanced BCa stage (T2 or greater) [21]. Moreover, data normalization to total protein instead to creatinine increases the sensitivity and specificity of these makers because total protein is less influenced by hematuria, which is commonly found in BCa patients [143]. Also, combining urine cytology with urine proteases profiling increases sensitivity from approximately 75% (given by conventional screening approaches) to more than 95%, and if the MMP-9/TIMP-2 or MMP-2/TIMP-2 ratios are used, a sensitivity close 100% might be attained [50].

Serine Proteases

Serine proteases represent more than one third of all identified proteases and include 40 families of well-known proteases, like trypsin and thrombin [144, 145]. This family of enzymes, which have a serine in the catalytic site, are involved in several biological processes from food digestion to inflammation response [145]. Thrombin-like proteases are a class of serine proteases involved in blood coagulation and fibrinolysis and include thrombin, plasminogen activators and plasmin [145]. Urokinase-type plasminogen activator (PLAU or uPA) is a serine protease that in addition to regulate fibrinolysis, also modulates innate and adaptive immune responses (Table 3, Figure 2), acting as an endogenous antibiotic (e.g. bacteriostatic against *Staphylococcus aureus*) [146], and leading to severe compromised T cell activation and proliferation if absent [147]. Its antimicrobial activity is mediated by the serine protease domain [148]. Moreover, PLAU is not only a marker of severe infection, but also a good predictor of mortality (76% sensitivity and 69% specificity for fatal sepsis) [149].

Despite the weak proteolytic activity, conversion of plasminogen to plasmin by uPA makes it a molecule of high interest for the study of tumors, since plasmin is related with various malignant properties of cancers [150]. In fact, the urokinase plasminogen

activator system, which comprises PLAU, is involved in several steps of cancer development and metastization (Table 3, Figures 1 and 2), including ECM degradation, cell proliferation, migration and adhesion, angiogenesis and intravasation [151–153]. PLAU can activate the precursor forms of MMP-9 and other matrix metalloproteases (like MMP-3 and MMP-12), leading to ECM remodeling. ECM remodeling is essential to angiogenesis and consequent tumor growth [151]. However, a study in human prostate cancer cells shows that PLAU may have an antiangiogenic role, due to antiangiogenic activity of angiostatin, produced by PLAU/plasmin system [154]. While there are two types of plasminogen activator (PLAU and tissue-type plasminogen activator), only PLAU promotes cell migration and proliferation via its growth factor domain [155]. In contrast, large reductions in the levels of its receptor (approximately 70%) has been shown to induce a dormancy state in human epidermoid-carcinoma (HEp3) cell lines [156]. PLAU also induces cell migration, via its kringle domain, by acting as a chemotactic agent [157]. Therefore, a continuum comprising infection-induced PLAU upregulation and PLAU-induced cellular proliferation/chemotaxis can be discerned. In BCa, PLAU modulates many signaling pathways, including NK-kappa B, TWEAK signaling pathway and Wnt signaling pathway, as highlighted in Table 3 and Figure 2. Furthermore, PLAU levels was suggested as a parameter for BCa prognosis. In fact, evidences show that patients with BCa displaying higher PLAU levels have lower survival rates than patients with low levels of PLAU [158].

The added value of Urine proteases for Bladder Cancer diagnosis

In order to better integrate the role of proteases in the pathogenesis of BCa (Figure 1) together with its clinical value, Table 3 overviews the differential involvement of urinary proteases in BCa development and progression. As can be depicted from this figure, MMP-2 and -9, PLAU and cathepsins D and G seem to be involved in all stages of BCa, including tumorigenesis and angiogenesis, cancer progression, invasion and metastasis. Therefore, the detection of these 4 proteases would most likely be indicative of BCa, independently of its stage. Moreover, these proteases might be seen as therapeutic targets in the set of BCa. However, the majority of clinical trials with MMP-inhibitors has failed, most likely because these were introduced only in advanced stages of the disease

[127]. In the search of more potent and specific agents, near-infrared polymer-based proteolytic beacons were tested in animal models and were able of detecting tumors as small as 0.01 cm^2. Techniques employing visible or near infrared fluorescence resonance energy transfer fluorophores to detect and measure MMPs' proteolytic activity allowed the detection of tumors at early stages with increased sensitivity [159]. Packard and colleagues have demonstrated that using a high-resolution fluorescent biosensor based on MMP (2, 9 or 14)-mediated peptide cleavage and fluorescence emission, the protease activity can be localized at the polarized leading edge of migrating tumor cells (rather than further back on the cell body) [160]. Similarly, biocompatible fluorogenic MMP substrates allow assessing the efficacy of MMP-inhibitors in tumors without the requirement for tissue biopsies [161]. Using this approach, MMP inhibition has been shown to take place within hours after treatment initiation using a potent MMP inhibitor as prinomastat [161]. Furthermore, radiolabeled imaging techniques such as positron emission tomography (PET) and single photon emission computed tomography (SPECT) may be employed for accessing MMPs' activity in order to screen patients and diagnose cancer [162, 163]. Nevertheless, as biomarkers assessed in biological fluids, MMPs and corresponding inhibitors are rather unspecific for Bca. So, envisioning to improve the clinical management of BCa, the most sensitive, specific and robust panels of urinary biomarkers must combine: urinary MMP-2 and MMP-9, their inhibitors and ratios thereof, cathepsins G, D, L and B, and PLAU (Table 3). However, one should keep in mind that these markers will definitively become more robust if combined with conventional markers of tumorigenesis and malignancy [164, 165].

Urinary proteases as therapeutic targets in the management of Bladder Cancer

Cathepsin, MMPs or PLAU are not primary therapeutic targets in the clinical management of BCa, despite the available broad-spectrum inhibitors particularly for MMPs (e.g. Batimastat, Marismastat, Salimatat, Prinomastat and Tanomastat). These MMPs inhibitors (MMPIs) have been tested for the treatment of cancers such as leukemias, lymphomas, testicular, lung, gastrointestinal, oropharangial cancer, once inhibit malignant growth by enhancing fibrosis around malignant lesions. By doing so,

MMPIs prevent tumor invasion, apoptosis and angiogenesis. However, these therapies induce significant side defects, which limits its clinical applicability [166].

Despite not directly targeting proteases activity, most of BCa therapies have impact on the regulation of proteases' expression and activity. For instance, tyrosine kinase inhibitors were reported to modulate proteases in the set of bladder cancer. For example, the treatment of bladder cancer cell lines with sunitinib impaired cathepsin B activation and stimulated a lysosomal-dependent necrosis, whereas pazopanib induced cathepsin B activation and autophagic cell death [167]. Glucocorticoids also impact the activity of proteases in BCa. Corticosterone, prednisone, and dexamethasone were shown to inhibit the expression of MMP-9 in bladder cancer cell lines with impact on cell invasion but marginal effect on cell growth [168].

Intravesical treatment with Bacillus Calmette- Guerin (BCG) is a clinically established and effective therapy for superficial bladder cancer and CIS though the mechanism of BCG immunotherapy is largely unknown. BCG was suggested to bind to fibronectin, near the carboxyl terminal region and adjacent to the heparin-binding domain, and protect this region of the molecule from tumor proteases [169]. More recently, the antitumor effect of BCG on bladder cancer was associated to BCG-induced cytokines TNF-α and INF-γ. Curiously, these cytokines were shown to induce MMP-9 expression and increase its enzyme activity in J-82 bladder cancer cells, potentially enhancing the invasiveness of bladder cancer in certain conditions [170]. Cathepsin B was also shown to mediate the BCG effect. The BCG-induced increase of this protease content and activity was related to the apoptosis of T24 and MB49 cell lines [171].

Immunotherapies with antibodies targeting PD-1 receptor expressed on T cells and its ligands, PD-L1 and PD-L2, have been recently proposed for the clinical management of BCa. Specifically, atezolizumab attaches to PD-L1 on the surface of tumor cells and prevents it from interacting with PD-1 receptors on immune cells and thus unleashes the immune system (ASCO 2016). These therapies also have impact on proteases activity once target MMP-dependent cleavage of PD-1 ligands on fibroblasts, inhibiting inflammation in tissues [172]. Nevertheless, few experimental evidences exist on this topic once these immunotherapies are starting to be implemented for the treatment of BCa.

In vivo evidences of the impact of these therapies on the levels and activity of these proteases in urine will help to disclose the biomarker value of urinary proteases for monitoring patients' response to therapeutics.

CONCLUSIONS

Proteases hyper activation is a hallmark of BCa, leading to tissue damage and invasion, immune system and apoptosis evasion, tumorigenesis and cancer metastasis. So, changes in the levels of proteases in biofluids as urine might help in the management of this malignancy. However, one should be aware that alterations in the gene expression patterns of proteases do not accurately reflect enzymes levels and activity. MS is particular suitable for the detection of quantitative and qualitative alterations of proteases in urine samples during exploratory research, as it provides an integrative view of the multiple molecular mediators, promoters and inhibitors that play a key role in BCa. In addition, zymography and immunoaffinity-based approaches are required for activity assays and validation of proteases as biomarkers, respectively. The application of these methodological approaches has shown the involvement of several proteases in tumor development, progression and dissemination. Nevertheless, these proteases have not been therapeutically targeted with success, but novel sensors are being developed and will hopefully allow the successful monitoring of urine proteases and therapy efficacy.

In conclusion, a fingerprint of proteolytic activity can be discerned in the urine of BCa patients, one comprising alterations in MMP-2 and MMP-9 activity, their inhibitors and ratios thereof, cathepsins G, D, L and B, DPEP and PLAU. Despite promising, the exploration of these proteases as biomarkers for the clinical management of BCa is still in its infancy. Future studies associating urine proteases profile with BCa stage, patients' age and gender, and response to therapeutics are needed to support their diagnosis and prognosis value.

Acknowledgements

This work was supported by the Portuguese Foundation for Science and Technology (FCT), European Union, QREN, FEDER and COMPETE for funding the iBiMED, UnIC, QOPNA research units (project UID/BIM/04501/2013, UID/IC/00051/2013, UID/QUI/UI0062/2013).

Table 1 - Procedures employed in bladder cancer urine proteomics

Procedure	Distinguishing Features	References
1D/2D Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis	High-throughput technique (hundreds to thousands of proteins). Useful for highly abundant and soluble proteins, allowing the separation of complex mixtures. Very acid or basic proteins can be accurately resolved. Allows bands/spots excised for MS analysis. Lacks dynamic range, reproducibility and sensitivity. Bad for poorly abundant (enzyme inhibitors) or poorly soluble (membrane) proteins. Laboriously coupled to MS analyzers.	[29–31,34,36–38]
Capillary Electrophoresis	Allows the study of naturally occurring peptides. Small amounts of sample required (nL). Very high speed (high voltage allowed) and resolution without band broadening due to high surface area and heat dissipation. Easily coupled to MS analyzers	[55–58]
Zymography	Detects enzymatic activity. Allows for multiple enzymes/substrates monitoring simultaneously. Allows prior molecular weight- and isoelectric point-based protein resolution, tissue sample *in situ* zymography, and analyses using protease-activated fluorogenic probes that allow the monitoring of multiple MPPs/fluorophores and the mapping of MMPs' activity .	[48–53]
Chromatographic separations	High resolution techniques. Separation based on several characteristics (mass, volume, isoelectric point, hydrophobicity). Susceptible to ion suppression by the presence of salts. High retention (e.g. column-to-column) times variability. Easily coupled to MS analyzers.	[30,32,33,35–37]
Western blot and other immunological assays	Allows relative and absolute quantification of multiple proteins. High resolution and sensitivity. Targeted approach. High Costs. Targeted (misses much information). Does not separate isoforms. Requires previous knowledge of its targets.	[30,36–38]

Table 2 – Classes of enzymes identified in the urine of bladder cancer humans patients detected by mass spectrometry and corresponding enriched biological processes			
Type of Enzyme	Number	Enriched Biological Processes	Enzymes
Oxidoreductase	6	Prostaglandin metabolic process, Hypochlorous acid biosynthetic process, Daunorubicin metabolic process, Doxorubicin metabolic process, Polyketide metabolic process, Peroxisome fission	ACOX1, AKR1C2, CP, DHRS2, NOS2, MPO
Transferase	11	Peptidyl-pyroglutamic acid biosynthetic process, using glutaminyl-peptide cyclotransferase, Activation of phospholipase A2 activity by calcium-mediated signaling, UDP-glucosylation, Glutamate catabolic process to aspartate or 2- oxoglutarate, L-kynurenine metabolic process, Double-strand break repair via alternative nonhomologous end joining, Ectopic germ cell programmed cell death, Positive regulation of catenin import into nucleus	GOT2, CD38, EGFR, F13A1, GSTT1, ART3, PLK2, PRKDC, QPCT, TGM4, UGGT1
Isomerase	2	DNA topological change, Cyclooxygenase pathway	PTGDS, TOP1
Lyase	4	Positive regulation of neurotrophin production, Double-strand break repair via classical nonhomologous end joining, Regulation of organelle transport along microtubule, Positive regulation of inositol 1,4,5-trisphosphate-sensitive calcium-release channel activity, Protein heterotetramerization, DNA demethylation, DNA ligation	XRCC6, CA1, APEX1, ADCY2
Hydrolase	21	Regulation of humoral immune response, Positive regulation of epidermal growth factor receptor signaling pathway, Positive regulation of ERBB signaling pathway, Regulation of complement activation, Regulation of protein activation cascade, Antibiotic metabolic	ADAMTS7, CTSD, CPN1, CTSG, CD38, CFB, C2, DDX39B, MASP2, MMP2, MMP9, PGA5, PON1, ALPP, PTPRC, LTF, DPEP1, VNN2, APEX1, PLAU, XRCC6

			process, Positive regulation of vascular smooth muscle cell proliferation, Negative regulation of tumor necrosis factor (ligand) superfamily member 11 production, Positive regulation of B cell proliferation, Endodermal cell differentiation, Positive regulation of neurotrophin production, Biofilm formation, Response to lactam/staurosporine/fluoxetine, Regulation of smooth muscle cell-matrix adhesion, Negative regulation of cell cycle checkpoint, Negative regulation of DNA damage checkpoint, Negative regulation of cysteine-type endopeptidase activity involved in apoptotic signaling pathway			

Table 3 - Proteases detected in distinct protein and mRNA levels in the urine of bladder cancer human patients

Type of protease	Protein Name	Gene Name	Uniprot Accession Number	Protein Variation	mRNA Variation	Bladder Cancer-associated Biological Processes/Pathways/Responses
Membrane dipeptidase	Dipeptidase 1	DPEP1	P16444	↓	↓	Dipeptides Hydrolysis, Eicosanoid Synthesis
Lysine carboxypeptidase	Carboxypeptidase N catalytic chain	CPN1	P15169	↑	↓	Steroid Biosynthesis
Serine endopeptidase	Cathepsin G	CTSG	P08311	↑	↓	ACE Inhibitor Pathway/Renin-angiotensin system, Lysosomal Trafficking
	Complement C2	C2	P06681	↑	↑	Complement and Coagulation Cascades, Pertussis, *Staphylococcus aureus* infection
	Complement factor B	CFB	P00751	↑	-	Complement and Coagulation Cascades,

						Staphylococcus aureus infection
	Urokinase-type plasminogen activator	PLAU	P00749	↓	↑	NF-kappa B signaling pathway, Complement and Coagulation Cascades, Wnt Signaling Pathway, Osteopontin Signaling, Transcriptional misregulation in cancer, Proteoglycans in cancer
	Mannan-binding lectin serine protease 2	MASP2	O00187	↓	↓/↑	Complement and Coagulation Cascades, Staphylococcus aureus infection
Aspartic endopeptidase	Cathepsin D	CTSD	P07339	↑	↓/↑	Lysosomal Trafficking
	Pepsin A-5	PGA5	P0DJD9	↑	-	Protein digestion and absorption
Metalloendopeptidase	A disintegrin and metalloproteinase with thrombospondin motifs 7	ADAMTS7	Q9UKP4	↑	-	Cartilage oligomeric matrix protein degradation
	Matrix metalloproteinase-9	MMP9	P14780	↑	↑	TWEAK Signaling Pathway, Osteopontin Signaling, Angiogenesis, AGE/RAGE pathway, Matrix Metalloproteinases Signaling, Leukocyte transendothelial migration, Estrogen signaling pathway, Transcriptional misregulation in cancer
	Matrix metalloproteinase-2	MMP2	P08253	↑	↓	AGE/RAGE pathway, Matrix Metalloproteinases Signaling, Leukocyte transendothelial migration, GnRH signaling pathway, Estrogen signaling pathway

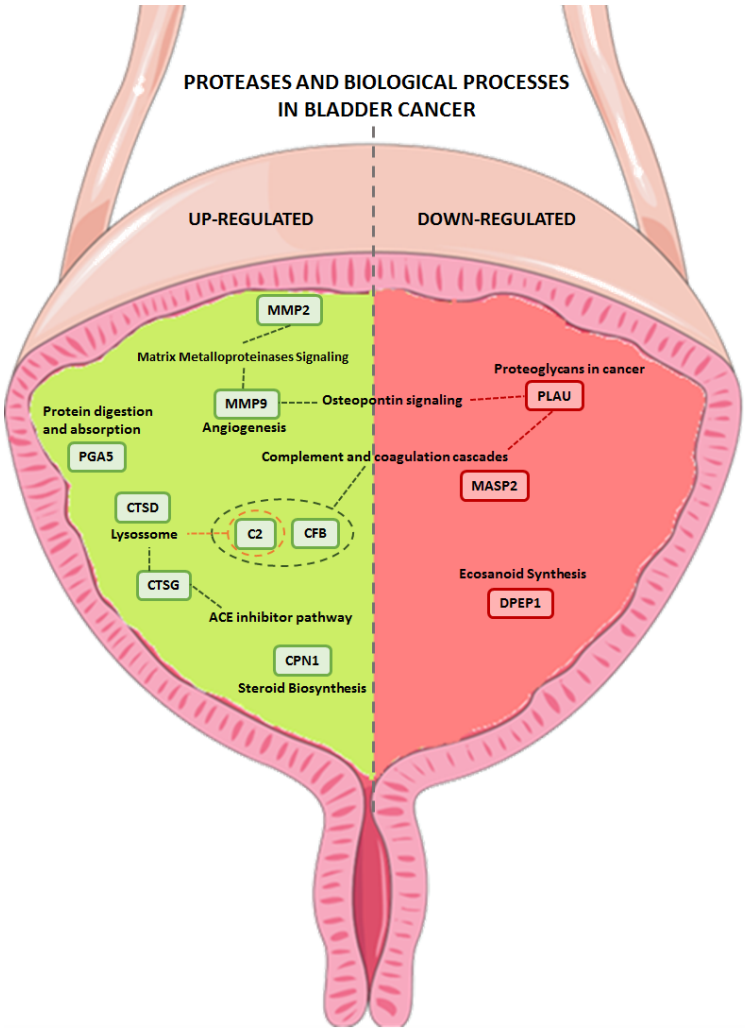


Figure 1 – Proteases modulated by bladder cancer and associated biological processes and putative biomarkers with diagnosis and prognosis value. Only a few representative examples are depicted. For a more comprehensive overview see Table 3. Figure was designed using *Servier Medical Art*.

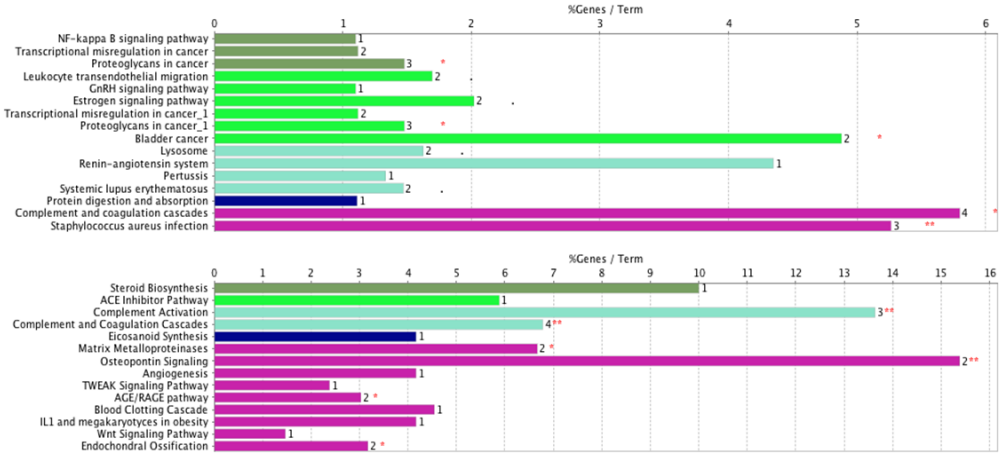


Figure 2 – Main pathways involving differentially expressed urinary proteases/peptidases modulated in Bladder Cancer human patients. Analysis performed using ClueGo plugin for Cytoscape. Upper Graph: KEGG pathways. Lower Graph: Wikipathways. Number after bars correspond to the number of genes contributing for those processes' enrichment. Bars marked with an asterisk (*) correspond to processes considered significantly changed (p-value <0.05) by the built-in algorithm.

References

1. Ferlay J, Soerjomataram I, Ervik M, et al (2014) GLOBOCAN 2012: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012: IARC CancerBase No. 11. In: Int. Agency Res. Cancer.

2. Fajkovic H, Halpern JA, Cha EK, et al (2011) Impact of gender on bladder cancer incidence, staging, and prognosis. World J Urol 29:457–463.

3. Madeb R, Messing EM (2004) Gender, racial and age differences in bladder cancer incidence and mortality. Urol Oncol Semin Orig Investig 22:86–92.

4. Bryan GT (1983) Pathogenesis of human urinary bladder cancer. Environ Health Perspect 49:201–207.

5. Lopez-Beltran A, Sauter G, Gasser T, et al (2004) World Health Organization classification of tumours: Pathology and genetics of tumours of the urinary system and male genital organs. Pathol. Genet. tumors Urin. Syst. male Genit. organs

6. WHO (2011) Environmental and occupational cancers. In: Fact sheet N°350.

7. Bethesda (2016) PDQ Adult Treatment Editorial Board. Bladder Cancer Treatment (PDQ®): Health Professional Version. In: Natl. Cancer Inst.

8. Yeung C, Dinh T, Lee J (2014) The Health Economics of Bladder Cancer: An Updated Review of the Published Literature. Pharmacoeconomics 32:1093–1104.
9. Johnson DC, Greene PS, Nielsen ME (2015) Surgical Advances in Bladder Cancer: At What Cost? Urol Clin North Am 42:235–252.
10. Mossanen M, Gore JL (2014) The burden of bladder cancer care: direct and indirect costs. Curr Opin Urol 24:487–491.
11. Svatek RS, Hollenbeck BK, Holmäng S, et al (2014) The Economics of Bladder Cancer: Costs and Considerations of Caring for This Disease. Eur Urol 66:253–262.
12. Rodrigues D, Jerónimo C, Henrique R, et al (2016) Biomarkers in bladder cancer: A metabolomic approach using in vitro and ex vivo model systems. Int J Cancer n/a–n/a.
13. Stenzl A, Cowan NC, De Santis M, et al (2011) Treatment of muscle-invasive and metastatic bladder cancer: update of the EAU guidelines. Eur Urol 59:1009–1018.
14. Kamat AM, Lamm DL (2004) Antitumor activity of common antibiotics against superficial bladder cancer. Urology 63:457–60.
15. Felix AS, Soliman AS, Khaled H, et al (2008) The changing patterns of bladder cancer in Egypt over the past 26 years. Cancer Causes Control 19:421–429.
16. Heyns CF, van der Merwe A (2008) Bladder cancer in Africa. Can J Urol 15:3899–3908.
17. Nagata M, Muto S, Horie S (2016) Molecular Biomarkers in Bladder Cancer: Novel Potential Indicators of Prognosis and Treatment Outcomes. Dis Markers 2016:8205836.
18. Kaufman DS, Shipley WU, Feldman AS (18AD) Bladder cancer. Lancet 374:239–249.
19. Massari F, Ciccarese C, Santoni M, et al (2016) Metabolic phenotype of bladder cancer. Cancer Treat Rev 45:46–57.
20. Lodillinsky C, Rodriguez V, Vauthay L, et al (2009) Novel Invasive Orthotopic Bladder Cancer Model With High Cathepsin B Activity Resembling Human Bladder Cancer. J Urol 182:749–755.
21. Gerhards S, Jung K, Koenig F, et al (2001) Excretion of matrix metalloproteinases 2 and 9 in urine is associated with a high stage and grade of bladder carcinoma. Urology 57:675–9.
22. Rosser CJ, Chang M, Dai Y, et al (2014) Urinary Protein Biomarker Panel for the Detection of Recurrent Bladder Cancer. Cancer Epidemiol Biomarkers Prev 23:1340–1345.
23. Lam T, Nabi G (2007) Potential of urinary biomarkers in early bladder cancer diagnosis. Expert Rev Anticancer Ther 7:1105–1115.
24. Yang N, Feng S, Shedden K, et al (2011) Urinary glycoprotein biomarker discovery for bladder cancer detection using LC/MS-MS and label-free quantification. Clin Cancer Res 17:3349–3359.
25. Urquidi V, Goodison S, Cai Y, et al (2012) A candidate molecular biomarker panel for the detection of bladder cancer. Cancer Epidemiol Biomarkers Prev 21:2149–2158.

26. Ye F, Wang L, Castillo-Martin M, et al (2014) Biomarkers for bladder cancer management: present and future. Am J Clin Exp Urol 2:1–14.
27. Burchardt M, Burchardt T, Shabsigh A, et al (2000) Current Concepts in Biomarker Technology for Bladder Cancers. Clin Chem 46:595–605.
28. Twining SS (1994) Regulation of Proteolytic Activity in Tissues. Crit Rev Biochem Mol Biol 29:315–383.
29. Li C, Li H, Zhang T, et al (2014) Discovery of Apo-A1 as a potential bladder cancer biomarker by urine proteomics and analysis. Biochem Biophys Res Commun 446:1047–52.
30. Roy R, Louis G, Loughlin KR, et al (2008) Tumor-specific urinary matrix metalloproteinase fingerprinting: Identification of high molecular weight urinary matrix metalloproteinase species. Clin Cancer Res 14:6610–6617.
31. Lei T, Zhao X, Jin S, et al (2013) Discovery of potential bladder cancer biomarkers by comparative urine proteomics and analysis. Clin Genitourin Cancer 11:56–62.
32. Chen CL, Lai YF, Tang P, et al (2012) Comparative and targeted proteomic analyses of urinary microparticles from bladder cancer and hernia patients. J Proteome Res 11:5611–5629.
33. Lindén M, Lind SB, Mayrhofer C, et al (2012) Proteomic analysis of urinary biomarker candidates for nonmuscle invasive bladder cancer. Proteomics 12:135–44.
34. Tsui K-H, Tang P, Lin C-Y, et al (2010) Bikunin loss in urine as useful marker for bladder carcinoma. J Urol 183:339–44.
35. Tan LB, Chen KT, Yuan YC, et al (2010) Identification of urine PLK2 as a marker of bladder tumors by proteomic analysis. World J Urol 28:117–122.
36. Feldman AS, Banyard J, Wu C-L, et al (2009) Cystatin B as a tissue and urinary biomarker of bladder cancer recurrence and disease progression. Clin Cancer Res 15:1024–31.
37. Smalley DM, Sheman NE, Nelson K, Theodorescu D (2008) Isolation and identification of potential urinary microparticle biomarkers of bladder cancer. J Proteome Res 7:2088–2096.
38. Lin C-Y, Tsui K-H, Yu C-C, et al (2006) Searching cell-secreted proteomes for potential urinary bladder tumor markers. Proteomics 6:4381–4389.
39. Staack A, Tolic D, Kristiansen G, et al (2004) Expression of cathepsins B, H, and L and their inhibitors as markers of transitional cell carcinoma of the bladder. Urology 63:1089–1094.
40. Svatek RS, Karam J, Karakiewicz PI, et al (2008) Role of urinary cathepsin B and L in the detection of bladder urothelial cell carcinoma. J Urol 179:478–84; discussion 484.
41. Ulrich F, Heisenberg CP (2009) Trafficking and cell migration. Traffic 10:811–818.
42. Malemud CJ (2006) Matrix metalloproteinases (MMPs) in health and disease: an overview. Front Biosci 11:1696.

43. Hadler-Olsen E, Fadnes B, Sylte I, et al (2011) Regulation of matrix metalloproteinase activity in health and disease. FEBS J 278:28–45.
44. Chen C-L, Lin T-S, Tsai C-H, et al (2013) Identification of potential bladder cancer markers in urine by abundant-protein depletion coupled with quantitative proteomics. J Proteomics 85:28–43.
45. Lee HJ, Nedelkov D, Corn RM (2006) Surface plasmon resonance imaging measurements of antibody arrays for the multiplexed detection of low molecular weight protein biomarkers. Anal Chem 78:6504–6510.
46. Ladd J, Taylor AD, Piliarik M, et al (2009) Label-free detection of cancer biomarker candidates using surface plasmon resonance imaging. Anal Bioanal Chem 393:1157–1163.
47. Gorodkiewicz E, Guszcz T, Roszkowska-Jakimiec W, Kozłowski R (2014) Cathepsin D serum and urine concentration in superficial and invasive transitional bladder cancer as determined by surface plasmon resonance imaging. Oncol Lett 8:1323–1327.
48. Mohammed MA, Seleim MF, Abdalla MS, et al (2013) Urinary high molecular weight matrix metalloproteinases as non-invasive biomarker for detection of bladder cancer. BMC Urol 13:25.
49. Õzdemir E, Kakehi Y, Okuno H, Yoshida O (1999) Role of matrix metalloproteinase-9 in the basement membrane destruction of superficial urothelial carcinomas. J Urol 161:1359–1363.
50. Eissa S, Ali-Labib R, Swellam M, et al (2007) Noninvasive Diagnosis of Bladder Cancer by Detection of Matrix Metalloproteinases (MMP-2 and MMP-9) and Their Inhibitor (TIMP-2) in Urine. Eur Urol 52:1388–1397.
51. Hattori S, Fujisaki H, Kiriyama T, et al (2002) Real-time zymography and reverse zymography: a method for detecting activities of matrix metalloproteinases and their inhibitors using FITC-labeled collagen and casein as substrates. Anal Biochem 301:27–34.
52. Vandooren J, Geurts N, Martens E, et al (2013) Zymography methods for visualizing hydrolytic enzymes. Nat Methods 10:211–20.
53. Moses MA, Wiederschain D, Loughlin KR, et al (1998) Increased incidence of matrix metalloproteinases in urine of cancer patients. Cancer Res 58:1395–1399.
54. Schönemeier B, Metzger J, Klein J, et al (2016) Urinary Peptide Analysis Differentiates Pancreatic Cancer From Chronic Pancreatitis. Pancreas.
55. Theodorescu D, Wittke S, Ross MM, et al (2006) Discovery and validation of new protein biomarkers for urothelial cancer: a prospective analysis. Lancet Oncol 7:230–240.
56. Jantos-Siwy J, Schiffer E, Brand K, et al (2009) Quantitative urinary proteome analysis for biomarker evaluation in chronic kidney disease. J Proteome Res 8:268–281.

57. Frantzi M, Van Kessel KE, Zwarthoff EC, et al (2016) Development and validation of urine-based peptide biomarker panels for detecting bladder cancer in a multi-center study. Clin Cancer Res.
58. Thongboonkerd V, Chutipongtanate S, Kanlaya R (2006) Systematic evaluation of sample preparation methods for gel-based human urinary proteomics: Quantity, quality, and variability. J Proteome Res 5:183–191.
59. Sedic M, Gethings LA, Vissers JPC, et al (2014) Label-free mass spectrometric profiling of urinary proteins and metabolites from paediatric idiopathic nephrotic syndrome. Biochem Biophys Res Commun 452:21–6.
60. Froehlich JW, Vaezzadeh AR, Kirchner M, et al (2014) An in-depth comparison of the male pediatric and adult urinary proteomes. Biochim Biophys Acta 1844:1044–50.
61. Valente MAE, Damman K, Dunselman PHJM, et al (2012) Urinary Proteins in Heart Failure. Prog Cardiovasc Dis 55:44–55.
62. Thongboonkerd V (2008) Urinary proteomics : towards biomarker discovery , diagnostics and prognostics. 810–815.
63. Lokeshwar VB, Habuchi T, Grossman HB, et al (2005) Bladder tumor markers beyond cytology: International Consensus Panel on bladder tumor markers. In: Urology. pp 35–63
64. Konety BR (2006) Molecular markers in bladder cancer: a critical appraisal. Urol Oncol 24:326–337.
65. Dancik GM (2015) An online tool for evaluating diagnostic and prognostic gene expression biomarkers in bladder cancer. BMC Urol 15:
66. Lindgren D, Sjödahl G, Lauss M, et al (2012) Integrated Genomic and Gene Expression Profiling Identifies Two Major Genomic Circuits in Urothelial Carcinoma. PLoS One 7:e38863.
67. Lindgren D, Frigyesi A, Gudjonsson S, et al (2010) Combined Gene Expression and Genomic Profiling Define Two Intrinsic Molecular Subtypes of Urothelial Carcinoma and Gene Signatures for Molecular Grading and Outcome. Cancer Res 70 :3463–3472.
68. Kim W-J, Kim E-J, Kim S-K, et al (2010) Predictive value of progression-related gene classifier in primary non-muscle invasive bladder cancer. Mol Cancer 9:3.
69. Riester M, Taylor JM, Feifer A, et al (2012) Combination of a Novel Gene Expression Signature with a Clinical Nomogram Improves the Prediction of Survival in High-Risk Bladder Cancer. Clin Cancer Res 18 :1323–1333.
70. Blaveri E, Simko JP, Korkola JE, et al (2005) Bladder Cancer Outcome and Subtype Classification by Gene Expression. Clin Cancer Res 11 :4044–4055.
71. Dyrskjøt L, Zieger K, Real FX, et al (2007) Gene Expression Signatures Predict Outcome in Non–Muscle-Invasive Bladder Carcinoma: A Multicenter Validation Study. Clin Cancer Res 13 :3545–3551.

72. Dyrskjøt L, Kruhøffer M, Thykjaer T, et al (2004) Gene Expression in the Urinary Bladder: A Common Carcinoma in Situ Gene Expression Signature Exists Disregarding Histopathological Classification . Cancer Res 64 :4040–4048.
73. Smith SC, Baras AS, Owens CR, et al (2012) Transcriptional Signatures of Ral GTPase Are Associated with Aggressive Clinicopathologic Characteristics in Human Cancer. Cancer Res 72 :3480–3491.
74. Sanchez-Carbayo M, Socci ND, Lozano J, et al (2006) Defining Molecular Profiles of Poor Outcome in Patients With Invasive Bladder Cancer Using Oligonucleotide Microarrays. J Clin Oncol 24 :778–789.
75. Choi W, Porten S, Kim S, et al (2014) Identification of distinct basal and luminal subtypes of muscle-invasive bladder cancer with different sensitivities to frontline chemotherapy. Cancer Cell 25:152–165.
76. Atapattu L, Lackmann M, Janes PW (2014) The role of proteases in regulating Eph/ephrin signaling. Cell Adh Migr 8:294–307.
77. Serim S, Haedke U, Verhelst SHL (2012) Activity-Based Probes for the Study of Proteases: Recent Advances and Developments. ChemMedChem 7:1146–1159.
78. López-Otín C, Bond JS (2008) Proteases: Multifunctional Enzymes in Life and Disease. J Biol Chem 283 :30433–30437.
79. Felix K, Gaida MM (2016) Neutrophil-Derived Proteases in the Microenvironment of Pancreatic Cancer -Active Players in Tumor Progression. Int J Biol Sci 12:302–313.
80. Pranjol MZI, Gutowski N, Hannemann M, Whatmore J (2015) The Potential Role of the Proteases Cathepsin D and Cathepsin L in the Progression and Metastasis of Epithelial Ovarian Cancer. Biomolecules 5:3260–3279.
81. Olson OC, Joyce JA (2015) Cysteine cathepsin proteases: regulators of cancer progression and therapeutic response. Nat Rev Cancer 15:712–729.
82. Drag M, Salvesen GS (2010) Emerging principles in protease-based drug discovery. Nat Rev Drug Discov 9:690–701.
83. Turk B (2006) Targeting proteases : successes , failures and future prospects. 5:785–799.
84. Vandooren J, Opdenakker G, Loadman PM, Edwards DR (2016) Proteases in cancer drug delivery. Adv Drug Deliv Rev 97:144–155.
85. López-Otín C, Matrisian LM (2007) Emerging roles of proteases in tumour suppression. Nat Rev Cancer 7:800–808.
86. Gocheva V, Joyce JA (2007) Cysteine cathepsins and the cutting edge of cancer invasion. Cell Cycle 6:60–64.
87. Shi GP, Villadangos J a, Dranoff G, et al (1999) Cathepsin S required for normal MHC class II peptide loading and germinal center development. Immunity 10:197–206.

88. Bania J, Gatti E, Lelouard H, et al (2003) Human cathepsin S, but not cathepsin L, degrades efficiently MHC class II-associated invariant chain in nonprofessional APCs. Proc Natl Acad Sci U S A 100:6664–9.
89. Stypmann J, Gläser K, Roth W, et al (2002) Dilated cardiomyopathy in mice deficient for the lysosomal cysteine peptidase cathepsin L. Proc Natl Acad Sci U S A 99:6234–9.
90. Gocheva V, Zeng W, Ke D, et al (2006) Distinct roles for cysteine cathepsin genes in multistage tumorigenesis. Genes Dev 20:543–556.
91. Kuester D, Lippert H, Roessner A, Krueger S (2008) The cathepsin family and their role in colorectal cancer. Pathol Res Pract 204:491–500.
92. Stoeckle C, Sommandas V, Adamopoulou E, et al (2009) Cathepsin G is differentially expressed in primary human antigen-presenting cells. Cell Immunol 255:41–45.
93. Kargi HA, Campbell EJ, Kuhn 3rd C (1990) Elastase and cathepsin G of human monocytes: heterogeneity and subcellular localization to peroxidase-positive granules. J Histochem Cytochem 38:1179–1186.
94. Shimoda N, Fukazawa N, Nonomura K, Fairchild RL (2007) Cathepsin g is required for sustained inflammation and tissue injury after reperfusion of ischemic kidneys. Am J Pathol 170:930–40.
95. Reeves EP, Lu H, Jacobs HL, et al (2002) Killing activity of neutrophils is mediated through activation of proteases by K+ flux. Nature 416:291–297.
96. Tkalcevic J, Novelli M, Phylactides M, et al (2000) Impaired Immunity and Enhanced Resistance to Endotoxin in the Absence of Neutrophil Elastase and Cathepsin G. Immunity 12:201–210.
97. Owen CA, Campbell MA, Sannes PL, et al (1995) Cell surface-bound elastase and cathepsin G on human neutrophils: A novel, non-oxidative mechanism by which neutrophils focus and preserve catalytic activity of serine proteinases. J Cell Biol 131:775–789.
98. Brinkmann V, Reichard U, Goosmann C, et al (2004) Neutrophil extracellular traps kill bacteria. Science (80-) 303:1532–1535.
99. Polanowska J, Krokoszynska I, Czapinska H, et al (1998) Specificity of human cathepsin G. Biochim Biophys Acta 1386:189–198.
100. Raymond WW, Trivedi NN, Makarova A, et al (2010) How immune peptidases change specificity: cathepsin G gained tryptic function but lost efficiency during primate evolution. J Immunol 185:5360–8.
101. Adkison AM, Raptis SZ, Kelley DG, Pham CTN (2002) Dipeptidyl peptidase I activates neutrophil-derived serine proteases and regulates the development of acute experimental arthritis. J Clin Invest 109:363–71.
102. Sambrano GR, Huang W, Faruqi T, et al (2000) Cathepsin G activates protease-activated receptor-4 in human platelets. J Biol Chem 275:6819–6823.

103. Maison CM, Villiers CL, Colomb MG (1991) Proteolysis of C3 on U937 cell plasma membranes. Purification of cathepsin G. J Immunol 147:921–926.
104. Drag B, Petersen LC (1994) Activation of pro-urokinase by cathepsin G in the presence of glucosaminoglycans. Fibrinolysis 8:192–199.
105. Reilly CF, Tewksbury DA, Schechter NM, Travis J (1982) Rapid conversion of angiotensin I to angiotensin II by neutrophil and mast cell proteinases. J Biol Chem 257:8619–8622.
106. Klickstein LB, Kaempfer CE, Wintroub BU (1982) The granulocyte-angiotensin system. Angiotensin I-converting activity of cathepsin G. J Biol Chem 257:15042–15046.
107. Shamamian P, Schwartz JD, Pocock BJZ, et al (2001) Activation of progelatinase A (MMP-2) by neutrophil elastase, cathepsin G, and proteinase-3: A role for inflammatory cells in tumor invasion and angiogenesis. J Cell Physiol 189:197–206.
108. Benes P, Vetvicka V, Fusek M (2008) Cathepsin D-Many functions of one aspartic protease. Crit Rev Oncol Hematol 68:12–28.
109. Nicotra G, Castino R, Follo C, et al (2010) The dilemma: Does tissue expression of cathepsin D reflect tumor malignancy? the question: Does the assay truly mirror cathepsin D mis-function in the tumor? Cancer Biomarkers 7:47–64.
110. Hah YS, Noh HS, Ha JH, et al (2012) Cathepsin D inhibits oxidative stress-induced cell death via activation of autophagy in cancer cells. Cancer Lett 323:208–214.
111. Dian D, Vrekoussis T, Shabani N, et al (2012) Expression of cathepsin-D in primary breast cancer and corresponding local recurrence or metastasis: An immunohistochemical study. Anticancer Res 32:901–905.
112. Lentari I, Segas I, Kandiloros D (2002) The importance of cathepsin's-D tissular detection in laryngeal squamous cell carcinoma. Acta Otorhinolaryngol Belg 56:383–9.
113. Paksoy M, Hardal U, Caglar C (2011) Expression of Cathepsin D and E-Cadherin in primary laryngeal cancers correlation with neck lymph node involvement. J Cancer Res Clin Oncol 137:1371–1377.
114. Dumartin L, Whiteman HJ, Weeks ME, et al (2011) AGR2 is a novel surface antigen that promotes the dissemination of pancreatic cancer cells through regulation of cathepsins B and D. Cancer Res 71:7091–102.
115. Lösch A, Schindl M, Kohlberger P, et al (2004) Cathepsin D in ovarian cancer: Prognostic value and correlation with p53 expression and microvessel density. Gynecol Oncol 92:545–552.
116. González-Vela MC, Garijo MF, Fernández F, et al (1999) Cathepsin D in host stromal cells is associated with more highly vascular and aggressive invasive breast carcinoma. Histopathology 34:35–42.
117. Ohri SS, Vashishta A, Proctor M, et al (2008) The propeptide of cathepsin D increases proliferation, invasion and metastasis of breast cancer cells. Int J Oncol 32:491–498.

118. Vashishta A, Ohri SS, Proctor M, et al (2006) Role of activation peptide of procathepsin D in proliferation and invasion of lung cancer cells. Anticancer Res 26:4163–4170.
119. Szajda SD, Darewicz B, Kudelski J, et al (2005) [Cancer procoagulant and cathepsin D activity in blood serum in patients with bladder cancer]. Pol Merkur Lek 18:651–653.
120. Tokyol C, Köken T, Demirbas M, et al (2006) Expression of cathepsin D in bladder carcinoma: Correlation with pathological features and serum cystatin C levels. Tumori 92:230–235.
121. Jean D, Rousselet N, Frade R (2006) Expression of cathepsin L in human tumor cells is under the control of distinct regulatory mechanisms. Oncogene 25:1474–84.
122. Tan G-J, Peng Z-K, Lu J-P, Tang F-Q (2013) Cathepsins mediate tumor metastasis. World J Biol Chem 4:91–101.
123. Page-McCaw A, Ewald AJ, Werb Z (2007) Matrix metalloproteinases and the regulation of tissue remodelling. Nat Rev Mol Cell Biol 8:221–33.
124. Egeblad M, Werb Z (2002) New functions for the matrix metalloproteinases in cancer progression. Nat Rev Cancer 2:161–174.
125. SUN N, ZHAO Q, YE C, et al (2014) Role of Matrix Metalloproteinase-1 (MMP-1)/Protease-activated Receptor-1 (PAR-1) Signaling Pathway in the Cervical Cancer Invasion. J Reprod Contracept 25:18–25.
126. Sternlicht MD, Werb Z (2001) How matrix metalloproteinases regulate cell behavior. Annu Rev Cell Dev Biol 17:463–516.
127. Kessenbrock K, Plaks V, Werb Z (2010) Matrix Metalloproteinases: Regulators of the Tumor Microenvironment. Cell 141:52–67.
128. Vasala K, Turpeenniemi-Hujanen T (2007) Serum tissue inhibitor of metalloproteinase-2 (TIMP-2) and matrix metalloproteinase-2 in complex with the inhibitor (MMP-2:TIMP-2) as prognostic markers in bladder cancer. Clin Biochem 40:640–644.
129. Mitsiades N, Yu WH, Poulaki V, et al (2001) Matrix metalloproteinase-7-mediated cleavage of Fas ligand protects tumor cells from chemotherapeutic drug cytotoxicity. Cancer Res 61:577–581.
130. Kleiner DE, Stetler-Stevenson WG (1999) Matrix metalloproteinases and metastasis. Cancer Chemother Pharmacol 43 Suppl:S42–51.
131. Hanemaaijer R, Sier CFM, Visser H, et al (1999) MMP-9 activity in urine from patients with various tumors, as measured by a novel MMP activity assay using modified urokinase as a substrate. In: Ann. N. Y. Acad. Sci. pp 141–149
132. Margulies IM, Hoyhtya M, Evans C, et al (1992) Urinary Type-IV Collagenase - Elevated Levels Are Associated with Bladder Transitional Cell-Carcinoma. CANCER Epidemiol BIOMARKERS Prev Vol 1:467–474.

133. Cowden Dahl KD, Symowicz J, Ning Y, et al (2008) Matrix metalloproteinase 9 is a mediator of epidermal growth factor-dependent E-cadherin loss in ovarian carcinoma cells. Cancer Res 68:4606–4613.
134. Maretzky T, Reiss K, Ludwig A, et al (2005) ADAM10 mediates E-cadherin shedding and regulates epithelial cell-cell adhesion, migration, and beta-catenin translocation. Proc Natl Acad Sci U S A 102:9182–7.
135. Massagué J (2008) TGFbeta in Cancer. Cell 134:215–30.
136. Yu Q, Stamenkovic I (2000) Cell surface-localized matrix metalloproteinase-9 proteolytically activates TGF-beta and promotes tumor invasion and angiogenesis. Genes Dev 14:163–176.
137. Waldhauer I, Goehlsdorf D, Gieseke F, et al (2008) Tumor-associated MICA is shed by ADAM proteases. Cancer Res 68:6368–6376.
138. Le Maux Chansac B, Missé D, Richon C, et al (2008) Potentiation of NK cell-mediated cytotoxicity in human lung adenocarcinoma: Role of NKG2D-dependent pathway. Int Immunol 20:801–810.
139. Bergers G, Brekken R, McMahon G, et al (2000) Matrix metalloproteinase-9 triggers the angiogenic switch during carcinogenesis. Nat Cell Biol 2:737–44.
140. Olson MW, Bernardo MM, Pietila M, et al (2000) Characterization of the monomeric and dimeric forms of latent and active matrix metalloproteinase-9: Differential rates for activation by stromelysin 1. J Biol Chem 275:2661–2668.
141. Provatopoulou X, Gounaris A, Kalogera E, et al (2009) Circulating levels of matrix metalloproteinase-9 (MMP-9), neutrophil gelatinase-associated lipocalin (NGAL) and their complex MMP-9/NGAL in breast cancer disease. BMC Cancer 9:390.
142. Yan L, Borregaard N, Kjeldsen L, Moses MA (2001) The high molecular weight urinary matrix metalloproteinase (MMP) activity is a complex of gelatinase B/MMP-9 and neutrophil gelatinase-associated lipocalin (NGAL): Modulation of MMP-9 activity by NGAL. J Biol Chem 276:37258–37265.
143. Eissa S, Ahmed MI, Said H, et al (2004) Cell cycle regulators in bladder cancer: relationship to schistosomiasis. IUBMB Life 56:557–564.
144. Di Cera E (2009) Serine Proteases. IUBMB Life 61:510–515. doi: 10.1002/iub.186
145. Almonte AG, Sweatt JD (2011) Serine proteases, serine protease inhibitors, and protease-activated receptors: roles in synaptic function and behavior. Brain Res 1407:107–122.
146. Jin T, Bokarewa M, Tarkowski A (2005) Urokinase-type plasminogen activator, an endogenous antibiotic. J Infect Dis 192:429–437. doi: 10.1086/431600
147. Gyetko MR, Libre EA, Fuller JA, et al (1999) Urokinase is required for T lymphocyte proliferation and activation in vitro. J Lab Clin Med 133:274–288.
148. Vassalli JD (1985) A cellular binding site for the Mr 55,000 form of the human plasminogen activator, urokinase. J Cell Biol 100:86–92.

149. Uusitalo-Seppälä R, Huttunen R, Tarkka M, et al (2012) Soluble urokinase-type plasminogen activator receptor in patients with suspected infection in the emergency room: A prospective cohort study. J Intern Med 272:247–256.
150. Jankun J, Skrzypczak-Jankun E (1999) Molecular basis of specific inhibition of urokinase plasminogen activator by amiloride. Cancer Biochem Biophys 17:109–123.
151. Duffy MJ (2004) The Urokinase Plasminogen Activator System: Role in Malignancy. Curr Pharm Des 10:39–49.
152. Reuning U, Sperl S, Kopitz C, et al Urokinase-type Plasminogen Activator (uPA) and its Receptor (uPAR): Development of Antagonists of uPA / uPAR Interaction and their Effects In Vitro and In Vivo. Curr Pharm Des 9:1529–1543.
153. Andreasen PA, Kjøller L, Christensen L, Duffy MJ (1997) The urokinase-type plasminogen activator system in cancer metastasis: A review. Int J Cancer 72:1–22.
154. Gately S, Twardowski P, Stack MS, et al (1996) Human Prostate Carcinoma Cells Express Enzymatic Activity That Converts Human Plasminogen to the Angiogenesis Inhibitor, Angiostatin. Cancer Res 56 :4887–4890.
155. Rabbani S, Mazar A, Bernier S, et al (1992) Structural requirements for the growth factor activity of the amino- terminal domain of urokinase. J Biol Chem 267:14151–14156.
156. Aguirre Ghiso JA, Kovalski K, Ossowski L (1999) Tumor dormancy induced by downregulation of urokinase receptor in human carcinoma involves integrin and MAPK signaling. J Cell Biol 147:89–103.
157. Mukhina S, Stepanova V, Traktouev D, et al (2000) The chemotactic action of urokinase on smooth muscle cells is dependent on its kringle domain. Characterization of interactions and contribution to chemotaxis. J Biol Chem 275:16450–8.
158. Hasui Y, Marutsuka K, Suzumiya J, et al (1992) The content of urokinase-type plasminogen activator antigen as a prognostic factor in urinary bladder cancer. Int J Cancer 50:871–873.
159. McIntyre JO, Matrisian LM (2009) Optical proteolytic beacons for in vivo detection of matrix metalloproteinase activity. Methods Mol Biol 539:155–174.
160. Packard BZ, Artym V V., Komoriya A, Yamada KM (2009) Direct visualization of protease activity on cells migrating in three-dimensions. Matrix Biol 28:3–10.
161. Bremer C, Tung CH, Weissleder R (2001) In vivo molecular target assessment of matrix metalloproteinase inhibition. Nat Med 7:743–748.
162. Temma T, Sano K, Kuge Y, et al (2009) Development of a radiolabeled probe for detecting membrane type-1 matrix metalloproteinase on malignant tumors. Biol Pharm Bull 32:1272–1277.
163. Furumoto S, Takashima K, Kubota K, et al (2003) Tumor detection using 18F-labeled matrix metalloproteinase-2 inhibitor. Nucl Med Biol 30:119–125.
164. Smith SC, Theodorescu D (2009) The Ral GTPase pathway in metastatic bladder cancer: Key mediator and therapeutic target. Urol Oncol Semin Orig Investig 27:42–47.

165. Oxford G, Theodorescu D (2003) The role of Ras superfamily proteins in bladder cancer progression. J Urol 170:1987–93.
166. Chaudhary AK, Pandya S, Ghosh K, Nadkarni A (2013) Matrix metalloproteinase and its drug targets therapy in solid and hematological malignancies: An overview. Mutat Res - Rev Mutat Res 753:7–23.
167. Santoni M, Amantini C, Morelli MB, et al (2013) Pazopanib and sunitinib trigger autophagic and non-autophagic death of bladder tumour cells. Br J Cancer 109:1040–1050.
168. Ishiguro H, Kawahara T, Zheng Y, et al (2014) Differential regulation of bladder cancer growth by various glucocorticoids: corticosterone and prednisone inhibit cell invasion without promoting cell proliferation or reducing cisplatin cytotoxicity. Cancer Chemother Pharmacol 74:249–255.
169. D.L. C, W.P. S, J.C. C, et al (1994) Bacillus Calmette-Guerin interacts with the carboxyl-terminal heparin binding domain of fibronectin: implications for BCG-mediated antitumor activity. J Urol 152:1275–1280.
170. Belotti D, Paganoni P, Manenti L, et al (2003) Matrix metalloproteinases (MMP9 and MMP2) induce the release of vascular endothelial growth factor (VEGF) by ovarian carcinoma cells: Implications for ascites formation. Cancer Res 63:5224–5229.
171. Sandes E, Lodillinsky C, Cwirenbaum R, et al (2007) Cathepsin B is involved in the apoptosis intrinsic pathway induced by Bacillus Calmette-Guerin in transitional cancer cell lines. Int J Mol Med 20:823–828.
172. Dezutter-Dambuyant C, Durand I, Alberti L, et al (2016) A novel regulation of PD-1 ligands on mesenchymal stromal cells through MMP-mediated proteolytic cleavage. Oncoimmunology 5:e1091146.

C - BIOMARKER VALIDATION

THEORETICAL BACKGROUND

Single/Multiple Reaction Monitoring

The development of prognostic and diagnostic biomarker panels has been made possible with the advent of high-performance mass spectrometry (MS)-based technologies and approaches. In fact, the performance of even the most developed and optimized immunoassays has already been matched and surpassed by MS-based assays [1]. MS is an analytical technique based on the determination of mass-to-charge ratios (m/z) of analyte ions. In addition, as analytes with different masses or charges can display similar m/z ratios, MS-based techniques also rely on other properties in order to identify each ion, including its isotopic distribution and fragmentation patterns. Selected Reaction Monitoring (SRM) is a high-performance MS-based technique for precise targeted quantification, which relies on the detection and analysis of user-defined precursor and fragment ions [2]. Notably, this scanning mode has been proven highly reproducible, sensitive and robust over a wide dynamic quantification range and across many samples, being ideal for quantification purposes [3,4]. As a result, several applications of clinical relevance have been purposed for SRM. Among these, some have been exploited using urine samples, providing a good alternative for well-recognized biomarker quantification [5], a selective disease discrimination assay [6] and a tool for quantification of novel protein biomarkers. Also, when one considers fluids other than urine, SRM has been promisingly applied in the settings of Bladder Cancer (BCa) [7], lung cancer [8] and kidney injury [9], but many issues remain to be addressed [10].

Typically, protein biomarkers are selected based on their established or predicted involvement in a biological or disease-related process. As a first step, a target list of protein biomarkers to exploit is assembled based on previous experiments, scientific literature reports (see PeptideAtlas [11], Global Proteome Machine Database [12], Genome Annotating Proteomic Pipeline [13]) or prior in-house generated knowledge. From these point on, proteotypic peptides (specific for the precursor protein) have to be chosen [14,15]. Furthermore, for proteins not contained in databases, it is still possible to predict their corresponding proteotypic peptides using high-performance

algorithms, which use factors such as hydrophobicity, charge, energetic and structural properties to infer which proteotypic peptides are the most suitable ones. However, these tend to select different proteotypic peptides even if starting from the same protein set [16]. Because each protein can give rise to cleavage peptides with as much as 100-fold differences in intensity [17], the analysis process can be made more difficult when trying to combine target peptides spanning a huge concentration range. Still, combining different tools provides the best and acceptable starting points [11]. Nonetheless, predicted proteotypic peptides may still fail to get to validation phase due to several reasons, such as auto-degradation while queuing in the auto-sampler, lack of co-elution for the different fragment ions, contaminations, and others. Next, transitions (precursor-fragmentation peptide pairs) to be monitored have to be selected, tested, optimized, validated and refined. Herein, fragmentation-ions with the highest signal intensities, the lowest level of noise and the most reproducibility are the ones preferred. In the end, sequence assignments are subjected to statistical filtering and inference, so that only if an acceptable error rate is achieved a given peptide is deemed suitable (validated) for the intended purpose.

The Markers to be Validated

The following markers proposed for validation were selected based on prior preliminary exploratory results obtained at the Proteomics core facility of the Biomedical Research Foundation, Academy of Athens and were supported by other groups' findings.

SLIT2

The neuronal guidance factor Slit family of proteins is classically recognized for its role in the regulation of orientated neuronal cell growth and migration [18]. Moreover, the interaction of Slit and its 4 (known) Robo receptors is involved in the developmental processes of various organs. In addition, several studies also provide evidences that Slit/Robo signaling pathway plays an important role in tumor cell migration and angiogenesis [19]. A compromise of the interactions between Slit proteins with their Robo receptors can lead to suppression of the apoptosis process in malignant cells.

However, there are several Slit proteins (1-3) and Robo receptors (1-4), so that only a balance between all of these dictates the end result of alterations in this pathway.

Slit2–Robo4 interaction and activation of Robo4 receptor in colon cancer is known to inhibit the axon guidance protein netrin-1 [18]. In turn, netrin-1 prevents Deleted in Colorectal Cancer (DCC)-mediated apoptosis in cancer cells [20]. Therefore, downregulation of Robo4/Slit2 signaling enhances netrin-1 activity and blocks apoptosis [21]. Similarly, in both breast cancer and colorectal cancer, Slit/Robo pathway is known to inhibit cellular invasion by interaction with E-cadherin and β-catenin [22], so that downregulation of Slit/Robo pathway is responsible for local cellular invasion. In turn, in the case of liver cancer, Slit2/Robo1 is known to inhibit hepatocyte growth factor (HGF)-mediated cell migration [23]. Accordingly, Slit2/Robo1 inhibits chemotaxis and chemoinvasion [22], so that its suppression allows chemokine-dependent metastasis in various cell lines [24]. Moreover, the addition of recombinant Slit2 decreases chemokine-dependent metastasis [24,25], suggesting that there is room for pathway modulation in malignant diseases. Simultaneous, Slit2 and its receptors Robo1 and Robo4 inhibit chemokine-mediated leukocyte migration, which is in turn positively associated with poor cancer prognosis [26]. Hence, a decrease in Slit2 is expected to induce chemokine-mediated leukocyte migration and infiltration, local tissue damage and inflammation and cancer metastasis as a result of such immune cell infiltration [26].

Among the several cancer-related genes that are deregulated at the levels of DNA methylation and mRNA expression in BCa, one can find SLIT2. In particular, its promoter regions are hypermethylated (~60% cases) while its expression levels are drastically downregulated (~90%) among patients with bladder urothelial carcinoma [27]. Hence, downregulation of the anti-tumor SLIT2 gene expression levels allows tumor growth and cancer cell migration. As for its diagnostic and prognostic performance, SLIT2 methylation status has been evaluated together with two other markers of hypermethylation (HS3ST2 and SEPTIN9), in which case the 3 markers together were capable of identifying 91% of Ta, 100% of T1, 100% of CIS and 90% of other tumor stages, as well as 89% low-grade and 100% high-grade patients [28].

Because these expression patterns of Slit receptors follow the same ones observed for other types of cancer, in which case SLIT2 is known to be downregulated, SLIT2 protein

is expected to be downregulated in the same way as it is observed for other types of cancer. However, if this downregulation of SLIT2 protein is actually characteristic of BCa (or of a particular stage thereof) and if it can be monitored in urine remains to be explored.

PROF1

A significant increase of Profilin-1 (PROF1) has been reported in both urine and tissue samples collected from patients with invasive BCa as compared to non-invasive BCa and benign controls. However, high inter-individual variability is also the norm, which may result from sample factors other than PROF1 abundance, highlighting the limitations of the currently available assays. Irrespectively of this variability, an increase in PROF1 was previously shown strongly correlated with poor prognosis and increased mortality [29]. For these reasons, for the time being and considering the inter-individual variability, urinary PROF1 should be considered together with other markers for disease discrimination and stratification.

NMP22

The Nuclear Matrix Protein-22 (NMP22) is a marker of urothelial cell death that is found elevated in urine samples from patients with BCa [30]. Urinary NMP22 has been more frequently tested, with results far from ideal but promising nonetheless. In 1,331 patients with risk factors for BCa, NMP22 has been shown to achieve a sensitivity of 55.7% and specificity of 85% (as opposed to 15.8% and 99.2% for cytology) [31]. In recurrence cases, NMP22 was shown among 668 patients to achieve a sensitivity of 49.5% and specificity of 87.3% [32]. However, in another setting, NMP22 was shown to achieve a sensitivity of 65% and a specificity of 40% (in contrast with 44% and 78% for voided cytology, or 75% and 62% for washed cytology) [33]. Furthermore, immune-based NMP22 assays have been shown to achieve a sensitivity of 57%, but with a 19% false-positive rate, a lack of performance that may be at least partially explained by the high inter-individual variability observed across patients and institutions [34].

The major issue in BCa diagnosis concerns the sensitivity of the available tests rather than its specificity. While NMP22 has lower specificity and sensitivity compared to

cystoscopy, it is more sensitive and has a significantly lower cost than urine cytology [30]. Accordingly, while there seems to be some good diagnostic value for NMP22, its performance has to be enhanced by combining it with other biomarkers, particularly those providing higher sensitivities. Also, it should be taken into consideration that the available data originated from immunoassays, which are highly susceptible to sample matrix effects.

SPRC

Secreted protein acidic and rich in cysteine (SPARC or SPRC) is downregulated in BCa, an alteration though to allow bladder carcinogenesis and tumor cell progression and metastasis. Accordingly, a decrease in SPRC protein levels has been positively associated with dysplasia, neoplasia, metastasis and atypia [35]. However, this view is questioned by other studies, where both SPRC protein and SPARC gene have been reported as upregulated in bladder and other types of cancer as well. In particular, both high SPARC mRNA and SPRC protein levels in BCa tumors were shown to be associated with poor patient survival [36]. Therefore, these contradictory findings indicate that it will be a challenging task to interpret patterns of SPRC protein expression in BCa and to evaluate its diagnostic/prognostic performance. Also, factors other than quantitative alterations are suggested to account for these questionable findings, such as alterations in subcellular location or post-translational modifications, but there are no studies addressing these hypotheses.

VASP

Vasodilator-stimulated phosphoprotein (VASP) is among a small set of urinary BCa-specific exosomal proteins responsible for cellular architecture modulation, motility and adhesion, as well as tumorigenesis and metastasis. Since it is an exosomal protein, its concentration in urine is expected to be extremely low. Thus, the achievement of VASP accurate quantification may prove to be a challenging task. Irrespectively, VASP has been previously proven downregulated in invasive versus non-invasive BCa tumors by Western blot [37]. It is valuable to investigate the diagnostic performance of VASP urinary protein, as the secretion of exosomes and exosomal proteins is known to be

significantly altered in tumor cells undergoing active proliferation and invasion [38,39]. Accordingly, VASP is specifically involved in focal adhesion and trans-endothelial migration of leukocytes which promotes BCa invasiveness. Moreover, exosomes secreted by tumor cells arrange a microenvironment in the bladder that may cause phenotypic shifts of the normal epithelial cells, as these can act as recipients and capture tumor-derived exosomes [40]. Hence, it is valuable to investigate the diagnostic performance of VASP urinary protein in the setting of BCa.

Sample Preparation

The human proteome is highly complex and displays an enormous protein concentration dynamic range, which creates several limiting factors for protein and peptide-based biomarker discovery and validation. Not surprisingly if one considers the physiology of urine production, this complexity is also observed at the level of the urinary proteome. For one, it is extremely difficult if not impossible to quantify changes in abundance for several proteins simultaneously if their concentrations span several orders of magnitude. Second, there is an ideal amount of cleavage enzyme employed for each protein depending on its concentration, on the digestion efficiency and on the sample matrix itself. Accordingly, the amount of cleavage enzyme would ideally be optimized for each precursor protein, but when analyzing several ones with distinct endogenous abundances, a less than ideal compromise has to be reached. Third, absolute changes in low-abundance proteins may be more difficult to monitor than changes taking place in high-abundance ones and the later may mask the former [17].

Still, urine is the preferred sample when it comes to BCa biomarker discovery and validation, so that standardized protocols have been developed for handling it [41]. Urine samples from controls and BCa patients are usually collected in a clinical setting and frozen. After previously frozen samples are thawed, these are typically centrifuged to remove cells and other debris. At this point, sample preparation typically requires concentration, enrichment or isolation of all or a subset of urinary proteins, which can be achieved by protein precipitation (e.g. acetone, tricarboxylic acid, immunoaffinity), buffer exchange (e.g. dialysis), molecular weight/size centrifugation (e.g. exclusion size spin columns) or some sort of affinity-based extraction (e.g. affinity-based

chromatography). Then, total urinary protein concentration is estimated (e.g. Bradford or bicinchoninic acid assays) in order to normalize quantitative data and optimize downstream procedures, proteins are denatured (e.g. with urea buffer) in an appropriate buffer for subsequent steps (e.g. ammonium bicarbonate), reduced (e.g. dithioerythiol, dithiothreitol), alkylated (e.g. iodoacetamide) and digested (e.g. trypsin), at which point proteolytic peptides are further purified (e.g. C18 columns), dried (e.g. under vacuum) and resuspended in compatible buffer for chromatographic separation of proteolytic peptides [42,43]. After this point, peptide samples are separated by reverse-phase liquid chromatography (or its nano-flow counterpart, nano-HPLC) and eluted peptides are ionized during the passage from liquid to gas phase by a combination of applied voltage and drying sprayed gas (typically N_2) [44]. Then, once gas-phase ions are formed, samples are ready to be analyzed in a mass spectrometer (see next section).

Mass Spectrometric Analysis

Typically, prior to MS analysis, peptides are fractionated by LC. For reproducibility and optimization purposes, chromatographic conditions should be constant and peptide elution times should be normalized using a set of standards [45]. Accordingly, peptide elution times allow for acquisition scheduling and analysis during a relatively short time window when a peptide is expected to elute from the column (see below). With the addition of a standard synthetic peptide, its native counterpart is expected to co-elute and these are therefore amenable to comparisons. For this reason, it is possible to integrate the chromatographic peaks and to calculate relative quantities. Also, by knowing the absolute amount of standard injected into the column over a dilutions series, absolute quantification can be achieved for the endogenous peptides as well. Hence, one should always aim at using the most suitable ions for analysis when performing high performance targeted quantification.

As the best analytical performance has to be achieved, there is a preferential choice of Quadrupole-based analyzers for targeted MS-based quantification. Accordingly, Triple Quadrupoles (QQQ), Quadrupole/Time-of-Flight (QTOF) and Quadrupole/Ion-traps are the most frequently employed apparatuses due to their low noise, high transmission

efficiency (>95%) and rapid switch between transitions (~2 ms), which altogether provide high sensitivity and multiplexing capacities. In addition, this somewhat limited choice of apparatus helps with acquisition standardization and data comparison. Accordingly, optimal collision energies tend to be similar if the same type of apparatus is used across studies, which allows consistent detection of fragment ions.

Peptides elute at different times because they display distinct physicochemical properties. Together with the mass and charge characteristic of each ion, differences in elution times provide another characteristic property which helps to identify a specific peptide. In addition, differences in elution times allow for several precursor-fragment ion pairs (i.e. transitions) to be sequentially and repeatedly measured, therefore creating the possibility of concurrent quantification of multiple targets. When this is the case, the term Multiple Reaction Monitoring (MRM) MS is adopted instead of SRM MS [10].

As the number one choice, triple quadrupole (QQQ) mass spectrometers consist of 3 quadrupoles positioned in tandem (**Figure 1**). The 2^{nd} quadrupole acts as a collision cell where analyte ions are collided with a gas (e.g. Ar or N_2) and suffer fragmentation. Instead, the 2 other quadrupoles (1^{st} Q and 3^{rd} Q) act as mass filters and selectively monitor specific/specified molecular precursor ions (1^{st} Q) and 1 or several fragment ions (3^{rd} Q) generated from the precursor ion by collisional dissociation in the 2^{nd} Q. In the end, the number of ions reaching the final detector is proportional to and predictive of the amount of the initially selected analyte, so that an intensity trace can be created [46,47], **Figure 2**.

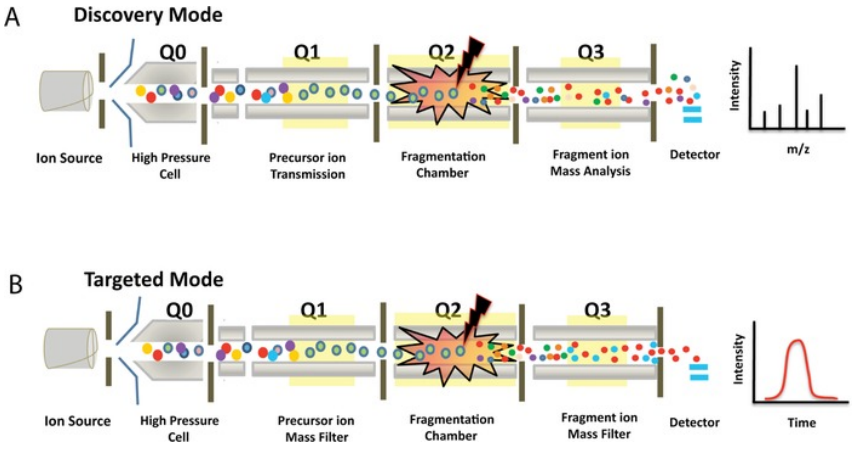


Figure 1 – Schematics depicting the main components of a triple quadrupole mass spectrometer and how it operates in A) discovery and B) targeted mode. In either case, Q1 serves as first filter for precursor ions, Q2 is the fragmentation chamber. In discovery mode, Q3 is turned off or channeled so that all product/fragmentation ions can pass through, allowing the most suitable peaks to be selected for targeted quantification. In turn, in targeted mode, Q3 acts as a second mass filter for product/fragmentation ions, allowing transitions (precursor-product ion pairs) to be analyzed [48].

Once in the collision cell (2nd Q), each peptide can generate a series of fragment ions: “to the left of”, “at the” or “to the right of” the peptide bound, meaning that these fragments can result from cleavage of the anomeric carbon-carboy carbon bound, the peptide bound, or the amine nitrogen-anomeric carbon bound, respectively. In each case, the charge can be located at the N- or C- terminus (the so-called a, b or c fragments if the charge is located in the N-terminus, or x, y or z fragments if the charge is located in the C-terminus of each of these bounds, respectively) (**Figure 3**). In addition, loss of water molecules and amine groups are also frequently observed, but transitions displaying it should be avoided because these tend to be inconsistent. In general, peptides prone to suffer chemical modifications (oxidation of W and M, deamination of N-G or Q-G, N-terminal cyclization of Q and E and carbamidomethylation of C) or other post-translational modifications should be avoided. Also, complete digestion should be ensured, and tryptic peptides with acidic residues (D, E), with two neighboring basic

amino acids at either cleavage site (KK, RR, KR, RK) or with N-terminal proline (KP, RP) should be avoided.

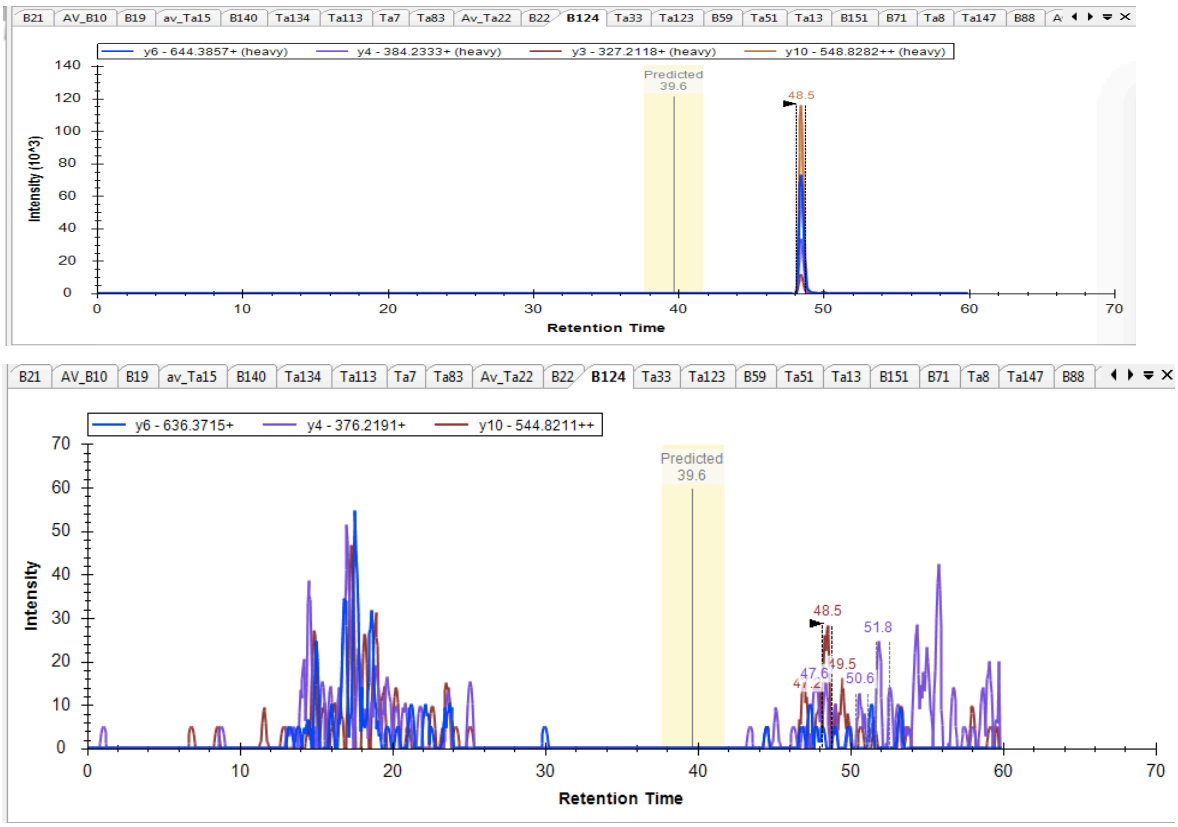


Figure 2 – Sample Extracted Ion Chromatogram from SLIT2 heavy (top) and light (bottom) peptides as viewed in Skyline, depicting chromatographic retention time on the x axis and signal intensity on the y axis. Total intensity for the peptide corresponds to the summed intensities of all transitions, herein depicted as different colored lines in the Extracted Ion Chromatogram. Note how i) not all transitions display the same intensity, ii) heavy and light counterparts display the same retention time, iii) the retention times differ from the predicted one, and iv) the light peptide elutes with other peptides displaying very flanking retention times when dealing with complex samples.

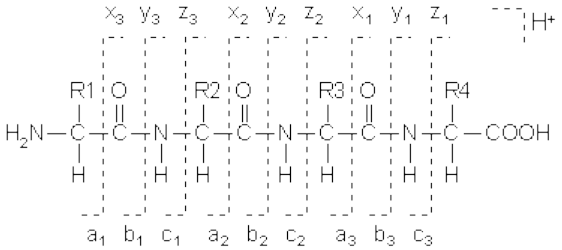


Figure 3 - Nomenclature for sequence ions in peptide mass spectra. Each peptide can generate a series of fragment ions, either "to the left of", "at the" or "to the right of" the peptide bound, meaning that these fragments can result from cleavage of the anomeric carbon-carboy carbon, the peptide, or the amine nitrogen-anomeric carbon bound, respectively. If the charge is located on the N-terminal fragment, the ion is classified as either a, b or c, depending on which bound is broken. If the charge is located on the C-terminal, the ion is classified as either x, y or z, depending on which bound is broken. A subscript indicates the number of residues in the fragment or the amino acid position along the peptide chain [49].

Peptide-based Biomarker Quantification

Clinical exploitation of a biomarker for the clinics requires the knowledge of its exact concentration so that it can be precisely compared with established reference concentration values (most commonly molarity or weight per volume). Currently, stable-isotope dilution (SID) is the most accurate, precise, reliable, reproducible and the gold standard method for absolute quantification. Isotopically labeled reference peptides or proteins (labeled with ^{13}C or ^{15}N, but not with ^{2}H because its chromatographic retention time is significantly affected) are chemically identical to their light native counterparts and therefore expected to behave the same way as their native light counterparts during sample handling, particularly when it comes to fragmentation [50]. As a major drawback, synthetic stable-isotope proteins [51] or peptides [50] are extremely expensive and not always possible to synthesize [52].

Major differences worth mentioning separate absolute from relative quantification. While the later presupposes comparisons between different groups and results on ratios of abundance, stable-isotope dilution-based absolute quantification presupposes the measurement of an intensity signal and its translation into a concentration value over a

range of concentrations. Therefore, precise quantification can only be achieved as long as an intensity signal from a sample of unknown amount falls within the linear range of the generated calibration curve. Also, it should be considered that each sample is unique and thus provides variable amounts of background, which affects chromatographic retention times and increases run-to-run and sample-to-sample variability [53,54] (the so-called “matrix effects”).

Several issues such as incomplete protein digestion, induced modifications during handling, unexpected native modifications, and partial loss of the reference synthetic peptide before addition, are to be expected to occur and should be considered when using synthetic stable-isotope proteins or peptides [55]. Nonetheless, it is possible to at least partially address them all by choosing the right peptides (the most obvious but less straightforward practice), practicing minimal handling, using protein standards instead of peptide standards and storing smaller aliquots so that fresh ones of synthetic peptide/protein can be more frequently used.

Alternatively, it is possible to heterologously express and metabolically label proteins composed of concatenated tryptic peptides, which can collectively be used as internal standards for different proteins or different proteotypic peptides from the same protein (QconCATs). However, a suitable expression system is required and such system can interfere with the efficiency of tryptic digestion, allowing the cleavage of QconCATs and endogenous proteins to be distinct and the different concatenated peptides to be cleaved to a different extent [56]. Nonetheless, this last issue can be addressed by tagging each peptide with a unique cleavable tag and normalizing the amount of tag released to the amount of expected peptide [57], but differences in cleavage efficiency between QconCATs and endogenous proteins in this case cannot be addressed because we are unable to control the cleavage process. Irrespectively, synthetic heavy peptides should be added to test samples as soon as possible (prior to digestion is regarded as a good rule of thumb) to minimize divergent results between native and heavy-labeled peptides derived from differential handling.

When absolute quantification is aimed, limits of detection (LODs) and limits of quantification (LOQs) become particularly critical because i) an assay as to be proven capable of consistently detect any given peptide considered for validation and ii) its

quantification limits (defined through the establishment of a calibration curve) have to include both the physiological and the pathological values found *in vivo*. While one could envision scenarios with relative quantification where values below or above the lower limits of detection or "quantification" could still be deemed valid and informative (if it is below/above these limits, it is still below/above the cut-off value, even if its accurate quantification is not possible), but same cannot be true when dealing with absolute quantification (if one wants to quantify a response to a drug or an analyte's concentration over time, precise quantification is required). In addition, matters are made more troublesome as these limits have to be established for each precursor-peptide and each transition (as intensities differ for each fragmentation-peptide). Once LODs and LOQs are established, one may still want to use and validate a biomarker whose abundance falls below or above one of these. If this is the case, it is possible to expand these limits by optimization of other MS parameters, including collision energy, decluttering potential or cone voltage. Again, these parameters are peptide-specific, so that optimization has again to be performed for each single analyte [58,59].

Ideally, multiple transitions *per* both heavy and light (native) peptides should be measured and multiple peptides *per* protein should be used, but due to technical (e.g. co-elution, lack of suitable fragment ions) and financial limitations (high cost of synthetic heavy peptides), this is frequently not achievable. Accordingly, 3 to 5 transitions for both heavy and light peptides are a realistic goal, one that can not unambiguously identify each protein from complex proteomes on its own but that provides considerable confidence (within a 1 Da window for each measurement) while being practical and sustainable.

The SRM cycle time is the product of and depends on the overall number of transitions and the time spent acquiring each transition/intensity (the dwell time, e.g. ~250 ms). Therefore, if 4 transitions are used for identifying 1 peptide (e.g. precursor peptide to y5, y7, y8 and y10 fragments) each one with a dwell time of 250 ms, the overall cycle time would be equal to 1 s. Higher dwell times (the time spent acquiring each transition) result in higher signal-to-noise ratio and, thus, in lower LODs for that particular transition. Accordingly, too many transitions (or too long dwell times) result in long cycle times *per* run, reducing the number of runs and the number of points available for

constructing a chromatogram. For this reason, a compromise between dwell time and cycle time has to be reached (the number of transitions would be optimized if these two other parameters were themselves optimally balanced), in which case high signal-to-noise can be achieved without compromising a minimal number of runs. As a rule of thumb, dwell times between 5 and 100 ms and 8 to 10 data points under the peak trace (or cycle times from 1 to 4 seconds) are a good starting point for efficient assays. This balance is achieved by scheduled SRM acquisitions, whereby transitions for each peptide are acquired in a small time window centered around its peptide elution time. This allows shorter but optimal time windows to be chosen, reducing the overall cycle time (with a higher number of runs) without diminishing the dwell time (maintaining high signal-to-noise ratios) [60]. However, it requires accurate retention times to be reproducibly observed. On top of that, it is also possible to perform conditional acquisitions, whereby multiple transitions are acquired for the same peptide only if the most representative of them (established during assay development) surpasses an intensity threshold, allowing time to be invested in the most robust acquisitions and avoiding time to be devoted to unnecessary or less informative ones [61].

Data Analysis

Method Refinement and Data Analysis

Mass and charge values (the m/z ratio), chromatographic retention times, fragment ions nature/relative intensities and high signal-to-noise ratios are expected to allow for unambiguous identification of target peptides. Also, an assay should be robust enough as to provide reproducible retention/elution times, trace shapes and fragment ions nature/relative intensities (fragmentation pathways). Achieving all these requirements simultaneously presupposes a level of optimization and refinement that can only be attained by the aid of assay-specific scripting. Because SRM/MRM comprises i) multiple transitions for the same and for different peptides, ii) multiple peptides for the same and for different proteins, iii) measurements of heavy and light samples for each of these, and iv) lots of noise and interferences resulting from different biological backgrounds and technical variability, the whole process of mass spectrometric analysis relies heavily on automation tools such as Automated and Targeted Analysis with

Quantitative SRM (ATAQS) [62], Skyline [63], MRMer [64] or one of several apparatus-specific software platforms. During the analysis process, common approaches are employed to partially address some of the aforementioned limitations, from target selection to transition optimization and post acquisition data analysis, including averaging [3] or summing [65] signals within groups (different transitions for the same peptide, different peptides for the same protein), analysis of variance (ANOVA) (to detect sources of variability or drivers of changes in abundance between groups) [66], paired and modified Student's t-tests (to analyze matched/controlled changes) [3] or regression mixed-effects models (to derive significance from protein/peptide-abundance changes at an higher complexity level) [67].

Regarding relative quantification of differential protein expression, data normalization has to be performed before any comparisons have taken place. If equal amounts of heavy isotope-labeled peptides have been spiked-in across samples to be compared, one way through which this can be accomplished is by adjusting for the median of the logarithmic intensities obtained from such heavy isotope-labeled peptides used as internal controls. Moreover, it is advantageous to analyze the same sample in these different settings (external control), to compare mean intensities and variances for each sample between the different settings and to normalize the remaining analytes accordingly [41] when comparing different sample preparations or data across laboratories.

When dealing with large datasets and trying to infer statistical significance from these differential expression, third-party softwares are usually employed. Most notably, the R environmental system has become the most widely used tool in the bioinformatics world, comprising a base programming language (rooted on the S language) and hundreds of different packages which provide several tools for data analysis and statistical inference. One of the major advantages of using R and similar tools relates to their reproducibility and optimization. These require the user to perform the analysis via scripts, which in turn can be made available, faithfully re-run and built upon to and by the scientific community. Therefore, the use of script-based statistical/data analysis is highly recommended, as the whole process can be re-evaluated, reproduced and modulated at any point. Irrespectively of however overwhelming SRM/MRM data may

be, it is import to manually/visually inspect it. Because analyzing all data is in some cases impracticable, starting at a higher level is extremely advantageous and made easier by software tools such as Skyline, which allows for a more superficial data treatment prior to its exportation to other formats [63].

Lastly, the accuracy and suitability of the intended biomarkers have to be determined, which is most often performed using receiver operating characteristic (ROC) analysis, which consist of plotted true-positive rates (so-called sensitivity or percentage of disease patients testing positive) versus false-positive rates (or percentage of healthy subjects testing positive for the disease) [68]. By doing so, even independent assays generated in different settings can be compared and the ability to distinguish between patients and controls most accurately discerned. In this setting, ROC analyses should be performed upon establishment of a multimarker classifier in order to access its clinical performance, which can be generated with Support Vector Machines (SVM).

Support Vector Machine Classifier

Support Vector Machines were initially developed by Cortes and Vapnik (1995) for binary classification [69], i.e. for discriminating any two groups based on a set of inputted observed features. Nonetheless, most of what is herein presented is based on the descriptions provided by Bennett and Campbell in their concept-centric tutorial [70] and the practical approaches on how to train and develop SVMs using R packages [71,72]. Basically, a SVM modulates a hyperplane capable of optimal class attribution, group discrimination or feature prediction. Such hyperplane is delimited by 2 margins or boundaries (1 on each group's side), the points lying on such boundaries are called support vectors, and the optimal separating hyperplane is located in the middle of the margins. Note, however, that in many complex problems accurate discrimination cannot be achieved with a linear model, in which cases higher order hyperplanes (e.g. sigmoid, radial) are required. When this is the case, data points are projected into a higher-dimensional space where these become linearly separable, a projection achieved using kernel techniques. Therefore, the problem can be solved as if it required a linear solution even though behind the scenes it's actually a non-linear one (which has major advantages, both computational and at the level of model robustness).

For optimization purposes, non-linear SVMs deal with data in the form of dot products only. First, imagine two vectors consisting of two sets of values $x = [x_1, x_2, x_3, x_4]$ and $y = [y_1, y_2, y_3, y_4]$. When trying to fit a model explaining these two co-variables, what is actually considered is the dot product x_1y_1, x_2y_2, etc. Why? Because dot product values provide a measure of similarity, as the dot product returns the cosine of the angle between them. Hence, if they are "parallel", it means that these tend to co-occur or change together and the returned value will be closer to 1. Otherwise, it will tend to 0 (in extreme cases, if the values are independent, the vectors will be perpendicular). Note, however, that there are multiple ways (types of SVMs) through which this can be achieved. For instance, in a Polynomial learning machine, the dot product is explicitly known $(x^Tx_j+1)^p$. In turn, for a radial basis/Gaussian function, $\exp(1/(2\sigma^2)||x-x_j||^2)$ is calculated instead, but with a sigmoid (neural net activation) function, $\tanh(\kappa x \cdot y - \delta)$ is how the algorithm proceeds. In either case, the dot products are calculated, but remember that in some cases it is not explicit, because the function was mapped into a higher order (non-linear modeling). In practice, what is calculated is the dot product of a function of x, f(x), times a function of y ,(f(y)), being the function specific for each type of kernel (e.g. sigmoid, radial) [70–72].

An interesting feature of SVMs is that data points on the "wrong" side of the discriminant margin (or that simply do not follow the overall pattern) are weighted down to reduce their influence, which avoids them from biasing or masking the pattern to be unveiled. Also, SVMs are modular, so that even though each of each can only solve binary classification problems, multiple (sub)classifiers can be combined to achieve a more complex solution.

Because the resulting learning algorithm is a generalization algorithm, its accuracy can be estimated by how well it performs on a test set. In other words, its accuracy is reflected on how well the model generalizes from the training (where it should perform optimally) to the test sets, which raises to key remarks. For one, the training and test sets should be independent (non-overlapping), as to avoid overfitting (which would only result in poor performance once the model is applied to an independent test population). Also, the training set should be represented of the test set; otherwise, the model will not be fit for its purpose [70–72].

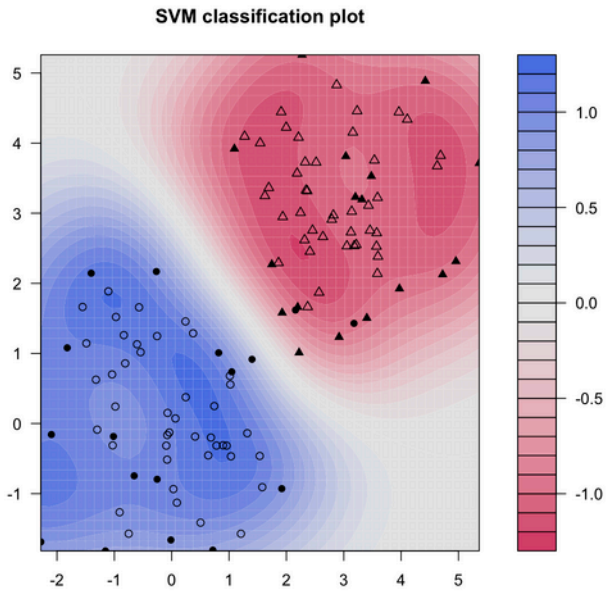


Figure 4 – Sample plot of a SVM classification model. Creating with a sample dataset using the kernlab R package.

As mentioned, in a SVM, the "support vectors" <are the data point closest to the decision plane/hyperplane, being the most difficult to classify but the most critical ones for defining the model. In fact, these points fully define the model, as the linear margins can only pass through a few such points. In what could become a limitation of SVMs, the points that define "optimality" are the ones in the critical region, where the differences are not so clear. Hence, poor analytical performance or faulty acquisition can severely compromise the model. In fact, in SVMs only these "support vector" points determine the weights and thus the boundaries and the hyperplane, so that changing/moving the extremes of the spectrum values (the data points further apart between each group) does not influence the boundaries and hyperplane definition. Therefore, the values less easily perceived as discriminating are the ones given more emphasis to. However, if relying on more extreme values, sensitivity is lost. We have previously mentioned that sensitivity is the major issue to be addressed in the setting of BCa. Hence, it does not matter how far the concentration spectra extends. In contrast, the important things lies in where we should draw the discriminating line. Therefore, although extreme values

may more accurately discriminate between BCa and healthy patients, they may have less of a role in defining a separating line between these.

SECTION II – EXPERIMENTAL METHODOLOGY

Methodology

Urine samples

Urine samples from controls and BCa patients were collected at the Urology clinical department of the Laikon University Hospital and Asklipieio General of Hospital of Voula, Athens, Greece, in accordance to the local ethical regulations. Summary clinical data on the urine samples are presented in **Table 1**. No major in differences were noted between disease and control patient groups regarding age, sex, urinary protein or urinary creatinine concentration. However, the proportion of patients with hematuria was twice as big in the control group (**Table 1**). For 8 samples, no complete laboratory parameters were available (other than disease group), so that these were excluded from the pooled clinical summary data in **Table 1**.

Table 1 – Summary Clinical Data

Group	Case	Control
Number	46 (71%)	19 (29%)
Cases Ta	14 (22%)	-
Cases T1	15 (23%)	-
Cases T2+	17 (26%)	-
Gender (Male/Female)	(40/6)	(16/3)
Age (Mean, StandDev)	67 (±11)	63 (±17)
Urinary Protein (mg/mL)	0.795 (±1.300)	0.595 (±0.945)
Urinary Creatinine (mg/dL)	108 (±53)	128 (±63)
Hematuria	8 (17%)	6 (32%)

Urinary Protein Extraction and Digestion

The underlying protein extraction procedure was based on the one proposed by Selevsek et al. 2011, with small adjustments [42]. After thawing, urine samples were centrifuged at 2,000 x g for 10 min at room temperature to remove cells and debris, and the urinary protein concentration was estimated by Bradford assay (Sigma-Aldrich, Saint Louis, MO). Due to sample handling restrictions and to assure precipitation and washing efficiency, the largest volume of sample precipitated at a time was 1 mL and always within a 500-1,000 µL range. For those samples requiring larger volumes to be

precipitated in order to achieve the minimum amount of protein per sample (30 µg for a 60 µL of 0.5 µg/µL final solution), multiple 1 mL aliquots were precipitated in separate. Because the maximum recovery achieved with the zip-tip cleaning step is 80 µg, urine samples with very large/high initially volume/concentration were precipitated in individual aliquots, as to avoid protein losses. Each sample volume was precipitated by overnight incubation at -20 °C on a 1:1 (v/v) ratio with trichloroacetic acid (15% TCA, final concentration 7.5% w/v) in the presence of sodium lauroyl sarcosinate (NLS, 0.1% v/v final concentration). Then, samples were centrifuged at 13,000 x g for 10 min at 4 °C, pellets were resuspended with 100 µL THF and centrifuged at 13,000 x g for 20 min at 4 °C to remove interfering compounds. Pellets were then dissolved in 100 µL of denaturing buffer (8 M urea in 0.05 M ammonium bicarbonate) and the total protein amount was determined by Bradford assay (Sigma-Aldrich, Saint Louis, MO). After precipitation, samples were resuspended in 100 µL denaturing buffer (8M urea in 0.05 M ammonium bicarbonate), from which a 10 µL aliquot was used to calculate protein concentration/protein yield by Bradford assay.

Based on protein yield, urinary proteins (in a 90 µL volume) were reduced (10 µL of a 0.1 M stock solution, final concentration 10 mM dithioerythiol, DTE) and alkylated (11 µL of a 0.5 M stock solution, final concentration 40 mM Iodoacetamide, IAA), diluted by adding more 700 µL of 0.05 M ammonium bicarbonate and digested by 16 hour-long incubations on a 1:100 (m/m) proportion of trypsin (0.5 µg/µL in advance prepared stocks) to sample protein (sequencing grade trypsin from Promega, Madison, WI). Afterwards, the digestion was terminated by adding formic acid (FA, 100% stock solution) to a final concentration of 0.1 %, peptides were desalted using Zip-tips (Thermo Scientific) and dried using a vacuum centrifuge. The final pellets were solubilized in appropriate volume of 0.1% formic acid (FA) (herein called buffer A) to a 0.5 µg/µL final concentration.

LC-MS setup

Liquid chromatography was performed using an Eksigent nanoLC-2D system from AB SCIEX coupled with a C18 nano-column (150 mm × 75 µm, particle size 5.0 µm) from Thermo Scientific. Peptide separation and elution was achieved with a 30 min 5-35%

ACN/water 0.1 % FA gradient at a flow rate of 300 nL/min. Six microliters of each sample were automatically collected with the autosampler and injected into the column. Urinary tryptic peptides were analyzed on an AB/MDS Sciex 4000 QTRAP with a nano-ESI source controlled by Analyst 1.5 software (Sciex). The mass spectrometer was operated in MRM mode, with the 1st (Q1) and 3rd quadrupole (Q3) at 0.7 unit mass resolution. Three-to-five transitions were recorded for the endogenous (light) and the internal standard (heavy) peptides. Optimum collision energies for each transition were automatically calculated by the Skyline software. Detailed information about the acquisition method and parameters used for each peptide are provided in **Table 2**.

Proteotypic Stable Isotope Labeled Synthetic Peptides Selection

Proteotypic peptides for SLIT2 (MW 1849.0 g/mol), SPRC (MW 1455.5 g/mol), PROF1 (MW 1222.3 g/mol), VASP (MW 1485.5 g/mol) and NMP22 (MW 1650.0 g/mol) were selected from PeptideAtlas (www.peptideatlas.org) [11] according to their score and number of observations in urine samples. Appropriate transitions for each peptide were selected with Skyline and the human spectral library (2014_05_29_human_consensus_final_true_lib) downloaded from NIST (National Institute of Standards and Technology, http://www.nist.gov/). One proteotypic peptide with 3-5 (top rank) transitions each were finally selected for analysis (**Table 2**). Crude synthetic isotopic labeled (^{15}N, and ^{13}C-labeled arginine or lysine) peptides were purchased from Thermo Scientific (Ulm, Germany) for method development and validation across samples. Heavy cysteine containing peptides were ordered alkylated with IAM to be identical to the urinary endogenous peptide analyzed by MRM.

Table 2 - Transition List with Acquisition Method for Detection and Quantification of each Proteotypic Peptide by MRM

Protein	Peptide Sequence	Precursor m/z	Product m/z	Fragment	Collision Energy (V)	Dwell Time (ms)
SLIT2_HUMAN	NHLQLFPELLFLGTAK (+2) light	614.348845	544.821149	y10	26.1	30
	NHLQLFPELLFLGTAK (+2) light	614.348845	319.197596	y3	26.1	30
	NHLQLFPELLFLGTAK (+2) light	614.348845	636.371538	y6	26.1	30

PROF1_HUMAN	NHLQLFPELLFLGTAK (+2) heavy	617.020244	548.828249	y10	26.1	30
	NHLQLFPELLFLGTAK (+2) heavy	617.020244	327.211795	y3	26.1	30
	NHLQLFPELLFLGTAK (+2) heavy	617.020244	644.385737	y6	26.1	30
	NHLQLFPELLFLGTAK (+2) heavy	617.020244	384.233259	y4	26.1	30
	DSPSVWAAVPGK (+2) light	607.31422	728.408986	y7	30.4	30
	DSPSVWAAVPGK (+2) light	607.31422	301.187031	y3	30.4	30
	DSPSVWAAVPGK (+2) light	607.31422	506.284734	y10 (+2)	30.4	30
	DSPSVWAAVPGK (+2) heavy	611.321319	736.423185	y7	30.4	30
	DSPSVWAAVPGK (+2) heavy	611.321319	309.20123	y3	30.4	30
	DSPSVWAAVPGK (+2) heavy	611.321319	510.291834	y10 (+2)	30.4	30
SPRC_HUMAN	TFDSSC[CAM]HFFATK (+3) light	483.21521	600.26113	y10 (+2)	22.1	30
	TFDSSC[CAM]HFFATK (+3) light	483.21521	542.747659	y9 (+2)	22.1	30
	TFDSSC[CAM]HFFATK (+3) light	483.21521	673.795337	y11 (+2)	22.1	30
	TFDSSC[CAM]HFFATK (+3) heavy	485.886609	604.26823	y10 (+2)	22.1	30
	TFDSSC[CAM]HFFATK (+3) heavy	485.886609	546.754758	y9 (+2)	22.1	30
	TFDSSC[CAM]HFFATK (+3) heavy	485.886609	677.802437	y11 (+2)	22.1	30
NMP22_HUMAN	QFC[CAM]STQAALQAMER (+3) light	547.58883	435.20203	y3	24.1	30
	QFC[CAM]STQAALQAMER (+3) light	547.58883	506.239143	y4	24.1	30
	QFC[CAM]STQAALQAMER (+3) light	547.58883	603.300786	y11	24.1	30
	QFC[CAM]STQAALQAMER (+3) heavy	550.924919	445.210299	y3	24.1	30
	QFC[CAM]STQAALQAMER (+3) heavy	550.924919	516.247412	y4	24.1	30
	QFC[CAM]STQAALQAMER (+3) heavy	550.924919	644.30599	y5	24.1	30
	QFC[CAM]STQAALQAMER (+3) heavy	550.924919	757.390054	y6	24.1	30
VASP_HUMAN	EEIIEAFVQELR (+2) light	738.390656	991.520721	y8	37.8	30
	EEIIEAFVQELR (+2) light	738.390656	791.441014	y6	37.8	30
	EEIIEAFVQELR (+2) light	738.390656	644.3726	y5	37.8	30
	EEIIEAFVQELR (+2) heavy	743.39479	1001.52899	y8	37.8	30
	EEIIEAFVQELR (+2) heavy	743.39479	872.486397	y7	37.8	30
	EEIIEAFVQELR (+2) heavy	743.39479	801.449283	y6	37.8	30
	EEIIEAFVQELR (+2) heavy	743.39479	654.380869	y5	37.8	30
	EEIIEAFVQELR (+2) heavy	743.39479	555.312455	y4	37.8	30

Proteotypic Peptide Detection and Method Development

In order to develop a method and to monitor the detection of the selected proteotypic peptides, a BCa sample (T2+_80) in which we had already identified SLIT2, PROF1, SPRC, VASP and NMP22 was used. The sample was prepared as described above and 1/20 of the final volume of the crude peptide mix was spiked in its tryptic digest. For these assays and clinical samples' analyses (note that for the quality controls wider ranges were tested), heavy peptides were used as follows: each 15 nmol (as provided in the acquired batches) were resuspended in 150 µL buffer A (initial concentration of 0.1 nmol/µL), all peptide were mixed in a final volume so that they were at a 1/50 of their initial concentration (typically 500 µL stock solutions containing 10 µL of each 0.1 nmol/µL heavy peptide were prepared) and these were used at a 1/20 of these 1/50 solutions (e.g. 8 µL for final volume of 160 µL).

Data Analysis and Protein Relative Quantification

SRM data were acquired on the QTRAP400 with Analyst 1.5 (AB Sciex). Peak detection and integration were determined on Skyline [63] based on two criteria: i) closest retention time and ii) highest correlation to the spectral library. All data were manually inspected with Skyline software to ensure correct peak detection, peak area selection and accurate integration. The peak areas for each peptide correspond to 3 µg (6 µL of 0.5 µg/µL solution injected for each) of total urinary protein.

Statistical Analysis for Classifier Generation

Exploratory and confirmatory basic statistical analyses for protein abundance and chromatogram areas were performed with the R base package. Boxplots were also created with the R base package, but ROC Curves were created on SPSS from IBM (Version 23.0).

Support vector machines (SVM) were created in order to combine the individual biomarker performances in a single classifier score. These multivariable-based classification models were generated using the R package "e1071". Three kernels were applied (linear, sigmoid and radial) for training the models, but only the best fit in each setting is presented. Arguments were optimized by running the algorithm multiple

times, each time changing these. The parameter cost was tested in the range of 2^{-10} to 2^{10} and parameter gamma in the range of $4^{0.01}$ to 4^{6}. Data was randomly divided into independent training (two-thirds) and test set (one-third) as to avoid overfitting. Therefore, each training dataset consisted of 13 control, 11 Ta, 10 T1 and 12 T2+ samples, and each test dataset consisted of 7 control, 6 Ta, 7 T1 and 7 T2+ samples.

SECTION III – EXPERIMENTAL RESULTS, DISCUSSION AND CONCLUSION

Results and Discussion

Automated Peak Quality Assignments

Even though Skyline provides tools for automated peak assignments and fast determination of how well the chromatographic runs were performed, these still require manual verification, correction and validation. One way of initially assessing the quality of the chromatographic runs was by looking at the retention times. As the same analyte and the exact same conditions applied to all samples, no major deviations should have been noted across samples. Similarly, the quality of peak identification and assignment was initially estimated by the dot product light-to-heavy (rdotp), a value between 0 and 1 which measures how well a peak matches its heavy counterpart. In order to display a good dot product light-to-heavy, a peptide must display the same retention time as its heavy counterpart and display the same transitions in the same proportion as its heavy counterpart. The higher these similarities, the higher the dot product light-to-heavy, which is typically deemed acceptable if above 0.80. As depicted in **Table 3**, the match between light and heavy peptides was identified by Skyline as excellent for the most part. However, the quality of these assignments tended to be overestimated because wrong peaks were sometimes automatically selected by Skyline. In turn, major deviations were observed for retention times across samples (**Table 3** and **Figure 5**) when it comes to PROF1, SPRC and NMP22 (minute-order deviations are too extreme, while the 0.21 min deviation in SLIT2 is acceptable).

While sometimes wrong, Skyline automated evaluations do come with some merit. In particular, they allowed the identification of patterns and outliers. For instance, it became evident early on which samples behaved the worst for NMP22 chromatographic resolution (T1_98, T2+110 and T2+122) (**Figure 5**). Similarly, it is notable that samples displaying higher retention times do so for all markers (e.g. Ta8 signal always lower between higher Ta83 and Ta7, the first five samples all increase in retention time and are followed by an abrupt decrease on the 6th one, samples B149, B151, B152 and B158 always appear to display the same pattern of retention times). This variability will become more evident and will be further discussed in the following sections, but it suggested from the beginning that the assumptions made about the promising

principles of “highly accurate” and “reproducible” LC-MS/MS analyses are far from straightforward to implement when dealing with urine samples.

Table 3 – Peptide Retention Time and Dot Product Light-to-Heavy

Parameter	SLIT2	PROF1	SPRC	NMP22
Peptide Retention Time, min (Mean ± St.Dev.)	48.42 (±0.21)	20.11 (±3.14)	18.66 (±3.19)	21.27 (±5.69)
DotProductLightToHeavy (Mean ± St.Dev.)	0.92 (±0.09)	0.98 (±0.04)	0.99 (±0.02)	0.93 (±0.09)

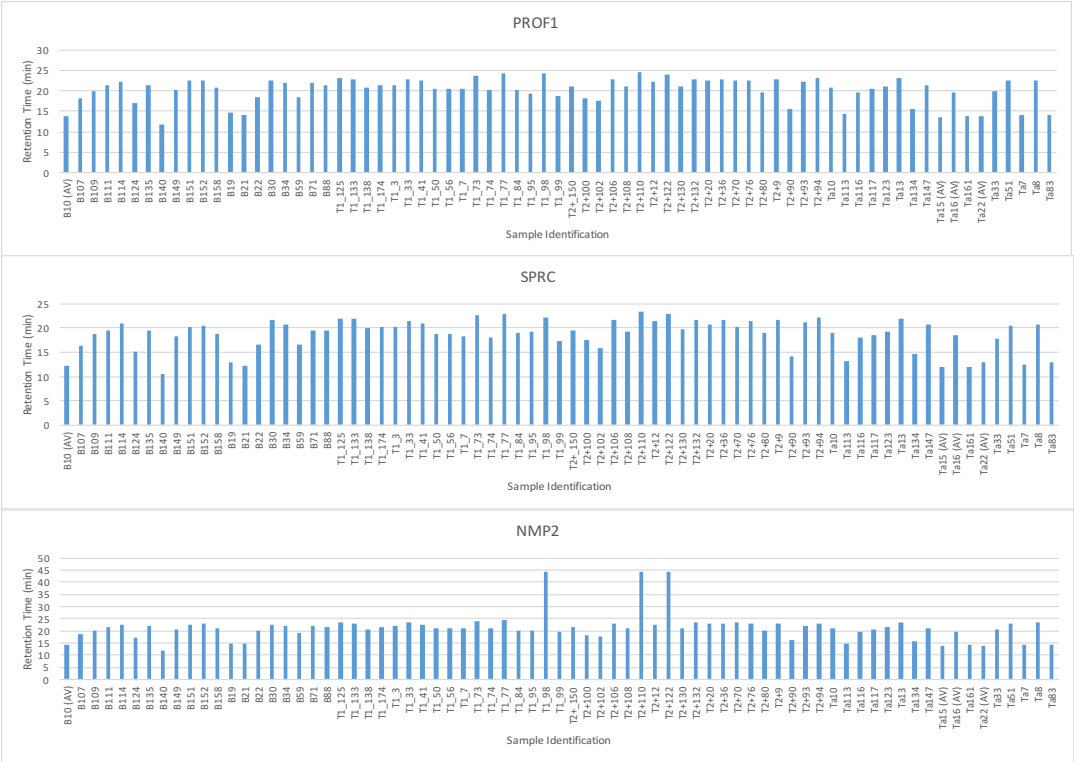


Figure 5 – Chromatographic retention times (in minutes) of PROF1, SPRC and NMP22 light peptides across samples. For SLIT2 retention time are not shown as they proved to be considerably stable across samples.

Quality Control Assays

Analytical assays are typically subject to method verification and validation for quality, reliability and robustness determination. Accordingly, an analytical method is only valid for the context in which it has been tested and validated, hence the term “suitability for intended purpose” (see [73] for compiled textbook-form detailed descriptions and [74,75] for international guidelines on method verification and validation). When analytical assay validation is aimed at, different components contributing for and dependent on data quality have to be considered: instrument qualification, method validation, system suitability and quality controls (QC) using QC standard samples. The herein presented work aimed at method development and dealt with qualification/verification challenges, but the same “validation” parameters were tested from the beginning as to allow assay optimization and to avoid unreliable investments. The same parameters used for assay validation can be employed to guide assay qualification/verification.

Instrumental qualification required the instrument specifications to be well defined, tested and confirmed so that the instruments were deemed suitable for the method. Then, qualified instruments (herein materialized by the HPLC and Triple-Q apparatuses) were combined with a specific analytical method (SRM/MRM) to run system-suitability tests. The SRM/MRM method here employed had been previously qualified in multiple studies and the suitability and analytical performance of the employed instruments were confirmed both prior to and during this project development by performing frequent QCs runs. However, the major challenge regarding its application for urine biomarker exploitation in the clinical setting (“suitability for intended purpose”) was not so straightforward to address. Nonetheless, we carried-out the task of evaluating this system’s suitability when dealing with urine samples.

QC Assays: Heavy Peptides Linearity in Buffer A

MRM analyses are expected to achieve high sensitivity (attomole level) and to provide absolute determination of peptide concentrations across a wide dynamic concentration range (covering up to 3-4 orders of magnitude, depending on multiple factors). In

addition, there are no “false-negatives” when it comes to SIS peptides, as these must be detected in every sample. Hence, if an endogenous peptide was not observed, it was considered as absent or to fall below the lower LOD for that particular assay. In these cases, peptides abundance was manually put to zero (0) in order to avoid background measurement. However, due to instrumentation availability limitations, the specific LODs over which the assay was performed could not be determined. Similarly, the signal intensity obtained through a peptide’s chromatogram is expected to be proportional to the amount of peptide injected. Hence, a linear relationship between peptide abundance and signal intensity over some orders of magnitude was to be expected. The concentration range over which peptide’s chromatogram intensity is expected to be linear is always narrower as compared to the one within which the peptide can be detected, as in the extremes of concentrations such linearity is naturally lost. However, as detailed below, such linearity was not assured.

Because the assurance of heavy standard peptides’ detection was a priority and would require less instrumentation time resources, the linearity of the heavy standard peptides’ chromatogram intensity for the concentrations to be used was assured by performing a dilution series in buffer A with concentrations flanking these target concentrations. Therefore, if observations of non-linearity took place, these would have to be attributed to urinary peptide sample constituents. For such purpose, solutions with higher, lower and equal concentrations of heavy peptides (compared to the ones used for clinical samples’ analysis) were injected, their intensity recorded and a calibration curve established. The signal was proportional to the amount of heavy peptide injected in buffer A (correlation > 0.95), which proved the peptides were chromatographically behaving linearly over a (at least) ten-fold concentration range in the same buffer as used for the HPLC sample runs (buffer A, 0.1% FA). Note that the signal intensity response obtained from the chromatograms of (only) heavy peptides injected at a concentration range from 20,000 (1/100 of neat stock) to 200,000 (1/10 of neat stock) attomole per µL for each heavy peptide, and that the clinical samples were instead analyzed together with 100,000 (1/20 of neat stock) attomole per µL of each heavy peptide, which falls within the linearity range. Hence, the qualified instruments

operating in the SRM/MRM scanning mode together with the heavy peptide analytes passed the system-suitability test.

QC Assays: Percentage Recovery and Analytical Performance

For method qualification and suitability for intended purpose determination, multiple parameters had to be assured. However, because this project was in its initial stages, only a few could be tested for. Assay linearity and accuracy, together with multimarker panel sensitivity and specificity were the chosen parameters requiring assement. Accuracy was defined as the extent to which the results generated by the method and the true values were in agreement.

The value for accuracy assessment could be obtained using a standard method. We assumed that how much of each peptide was being dealt with was known, as these came from commercial batches and no major sample handling steps had been performed. Therefore, because we were working with accurate targeted protein/peptide quantification, accuracy was assessed by analyzing the sample with known concentrations (from a certified reference) of heavy standards and comparing the measured intensity value to the one expected. This comparison of the measured response of the test sample as opposed to that of the reference material will be herein called “percentage recovery”. The concentration range over which Percentage Recovery must be determined should include either i) concentrations close to the quantitation limit, one in the middle of the range and one at the high end of the calibration curve, or ii) a critical decision value as the concentration point that must be the point of greatest accuracy. We opted by the second approach as LOQs had not been determined. Accordingly, the percentage recovery was tested by flanking the concentration range to-be-used in the clinical samples assay (1/20 of the 1/50 or 0.002 nmol/µL stock solution). Calibration points out of scale were considered non-reliable and excluded. Accordingly, 3 concentration groups (low, medium and high) were created for assessing the percentage recovery for each marker (1/20, 1/50, 1/100 for SLIT2; 1/10, 1/50, 1/100 for PROF1; 1/10, 1/20, 1/50 for SPRC and 1/10, 1/50, 1/100 for NMP22). The expected values (**Table 4**) were established by first creating a calibration curve using a dilution

series in buffer A for each marker individually. Then, a known amount of each peptide (in triplicate) for each concentration group (3x3 for each peptide) was injected and the average intensity values were used to calculate the initial amount of injected analyte using the calibration curve. Lastly, the true value of the injected amount and the predicted value using the calibration curve were compared, yielding the percentage recovery (**Table 4** and **Figure 6**).

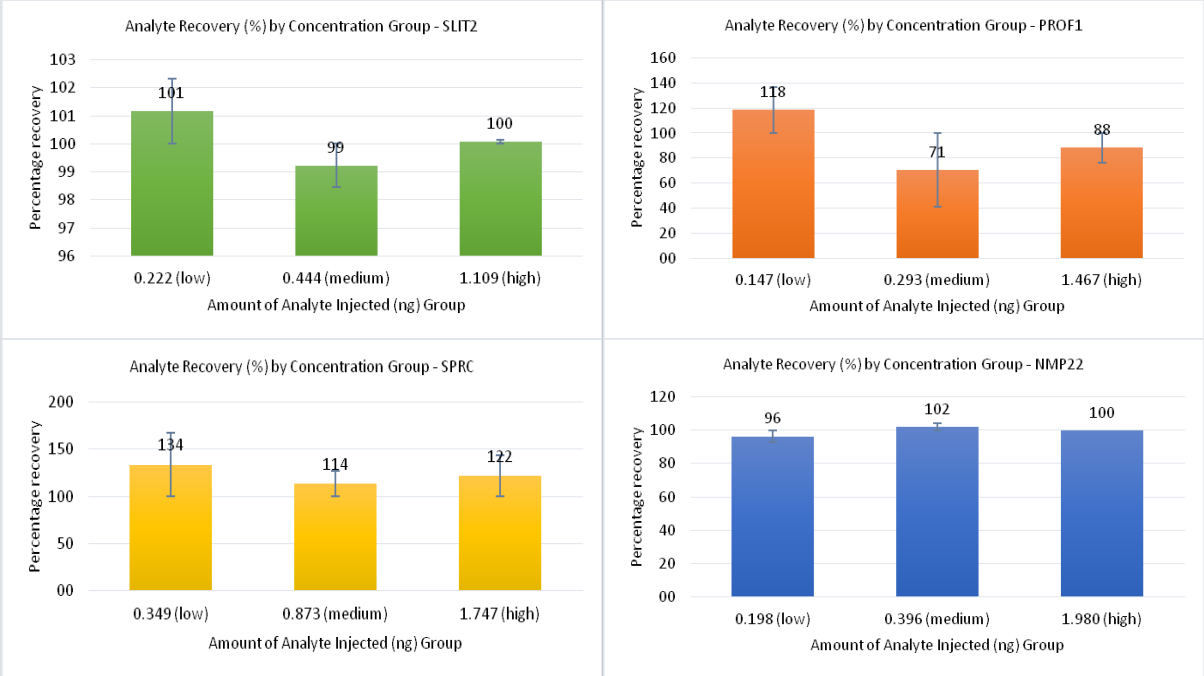


Figure 6 – Percentage Recovery plots for each analyte. These were established by first creating a calibration curve using a dilution series for each marker individually, injecting a known amount of each analyte in triplicate for each concentration group and comparing the intensity-based predicted amount to the known amount actually injected. The percentage recovery of VASP was not determined as initial data proved it would not be suitable for validation.

For assay qualification, accuracy (herein interpreted as percentage recovery) shall fall between 80% and 120%. While in some instances PROF1 (from 70.6% to 118.4% in one

case) and SPRC (from 113.6% to 133.7% in one case) did present with some considerable deviations, both these markers together with NMP22 (96.3% 102.1%) and SLIT2 (99.2% to 101.22%) presented near optimal recovery percentages. Note that despite the aforementioned deviations, these values were still better than or equal to those achieved with ELISA assays in urine and other not so problematic matrixes (see, for instance, a very recent work on ELISAs analytical performance in urine [76] or any of the multiple vendor-specific datasheets available on the web for other fluids such as plasma). For instance, available immunoassays for H2b have achieved an accuracy/percentage recovery from 64% to 298% and those performed for Survivin an accuracy/percentage recovery from 68% to 122%. Also, the percentage recovery herein achieved for SPARC (from 113.6% to 133.7%) was almost overlapping with that achieved with ELISAs performed in urine samples. Moreover, previous ELISAs have demonstrated an accuracy/percentage recovery from 50% to 69% for SLIT2, but we have herein achieved an accuracy from 99.2% to 101.2% when using MRM. Hence, both method and classifier seemed promising, but still required optimization. However, caution should be taken when interpreting such results, because these percentage recoveries were determined in buffer A and not in urine. As the sole purpose of these QC assays was to determine the accuracy of the analytical pipeline for these particular peptides (the system qualification) but not in the matrix itself, assay suitability in urine matrix had yet to be assessed.

Table 4 - Percentage Recovery Table for Heavy Standards

0.222 ng Spiked in Buffer A	mean SLIT2 (ng)	0.224
	expected SLIT2 (ng)	0.222
	% Recovery	101.2
0.444 ng Spiked in Buffer A	mean SLIT2 (ng)	0.440
	expected SLIT2 (ng)	0.444
	% Recovery	99.2
1.109 ng Spiked in Buffer A	mean SLIT2 (ng)	1.110
	expected SLIT2 (ng)	1.109
	% Recovery	100.1
0.147 ng Spiked in Buffer A	mean PROF1 (ng)	0.174
	expected PROF1 (ng)	0.147

	% Recovery	118.4
	mean PROF1 (ng)	0.207
0.293 ng Spiked in Buffer A	expected PROF1 (ng)	0.293
	% Recovery	70.6
	mean PROF1 (ng)	1.294
1.467 ng Spiked in Buffer A	expected PROF1 (ng)	1.467
	% Recovery	88.2
	mean SPRC (ng)	0.467
0.349 ng Spiked in Buffer A	expected SPRC (ng)	0.349
	% Recovery	133.7
	mean SPRC (ng)	0.992
0.873 ng Spiked in Buffer A	expected SPRC (ng)	0.873
	% Recovery	113.6
	mean SPRC (ng)	2.131
1.747 ng Spiked in Buffer A	expected SPRC (ng)	1.747
	% Recovery	122.0
	mean NMP22 (ng)	0.191
0.198 ng Spiked in Buffer A	expected NMP22 (ng)	0.198
	% Recovery	96.3
	mean NMP22 (ng)	0.404
0.396 ng Spiked in Buffer A	expected NMP22 (ng)	0.396
	% Recovery	102.1
	mean NMP22 (ng)	1.979
1.980 ng Spiked in Buffer A	expected NMP22 (ng)	1.980
	% Recovery	100.0

QC Assays: Constant amount of heavy peptides, variable amount of endogenous protein/sample matrix

For the same amount heavy peptide injected (2.5 µL of a 0.002 nmol/µL stock solution in a final volume of 50 µL, or herein also called 1/20 of the 1/50 stock solution), variable amount of endogenous total protein (from 1.5 to 6 µg) had a considerable matrix effect on the signal obtained from the heavy peptide (**Figure 7**). Note below that the analyses of the clinical samples were performed with the same amount of total endogenous

peptides employed across samples (at a final concentration of 0.5 µg/µL, or 3 µg). Thus, such matrix effects at the level of the heavy peptide-derived signal were excepted to be not so pronounced. However, this assumption presupposed the interference to result from endogenous proteins, which may not have been the case. Irrespectively, the samples were still considered for analysis, normalizing these to total amount of protein. Even with the observed interferences resulting from urine sample matrix, these findings still deserved much credit, as they revealed that the SRM/MRM assay was not robust enough to overcome the inherent matrix effects of urine samples. It is possible that chromatographic resolution was not high enough to discriminate between co-eluting/adjacent peaks typical of complex matrixes or that matrix effects resulting from interfering non-peptide metabolites caused artificial signal increase or suppression. Unfortunately, we were unable to discover the underlying cause of this interference phenomenon. These alterations of the heavy peptide-derived signal may have resulted from the presence of any of thousands of urinary species not fully removed during sample preparation. However, the nature of such interfering species was not amenable to discrimination. Irrespectively, the observed trend was clear, so that adding more endogenous peptides resulted in more pronounced heavy peptide signal deviations.

It was critical to ensure that the transitions on which the quantitation was based were free of interferences/contaminants from other peptides and components from the matrix. However, even though heavy peptide retention times were extremely consistent and reproducible from one sample to the other, the same could not be told of their intensity, so that interference-free quantification could not be ensured for any sample. In fact, even when looking at the heavy peptide-derived chromatograms only, some signal splits were noted. Such observation suggested the possibility of heavy peptides binding some other species (perhaps more highly abundant urinary peptides, to the column itself or to other urinary metabolites, particularly lipophilic substances). If this was the case, the heavy peptide-derived signal would be shifted, as both the molecular weight and the hydrophobicity of the “conjugate” would be different, and a “split” between its free and bound form would be observed. Therefore, the intensity of its free form (at the expected retention time) would be different (lower) than expected, and the shift would be more pronounced with higher amounts of urinary endogenous peptides

co-injected. As predicted, this phenomenon was indeed observed (**Figure 7**). In fact, the presence of higher amounts of endogenous tryptic peptides was accompanied by a linear decrease in the heavy peptide-derived intensity for each of these peptides (**Figure 6**). Note that "with higher amounts of urinary endogenous peptides" is referred to without discarding the possibility of urinary metabolites coexistence among the endogenous peptides. As the same amount of endogenous urinary peptides was obtained by starting with different urine volumes across samples (more diluted samples requiring larger volumes), variable amounts of other interfering substances may have been introduced. Therefore, higher levels of possible contaminants other than urinary peptides could be discarded. Supporting this hypothesis, was the fact that when splits did appear, the new signal would display a considerable tailing, characteristic of the presence of lipophilic substances. However, if this hypothesis was true, and if the whole assay were to require some sort of contaminant depletion, it also meant that the assay would most likely not make it into the clinics, compromising its sole purpose. Still, as above mentioned, the same amount of total endogenous peptides was employed across samples, which was expected to ameliorate this phenomenon (but which would probably leave the assay unreliable nonetheless). Again, such amelioration would only be expected if the matrix effects were due to other urinary peptides but not if resulting from other species. Nonetheless, we proceeded with the clinical samples' analysis as it would be possible to measure the heavy signals and determine if these remained constant over time and across samples.

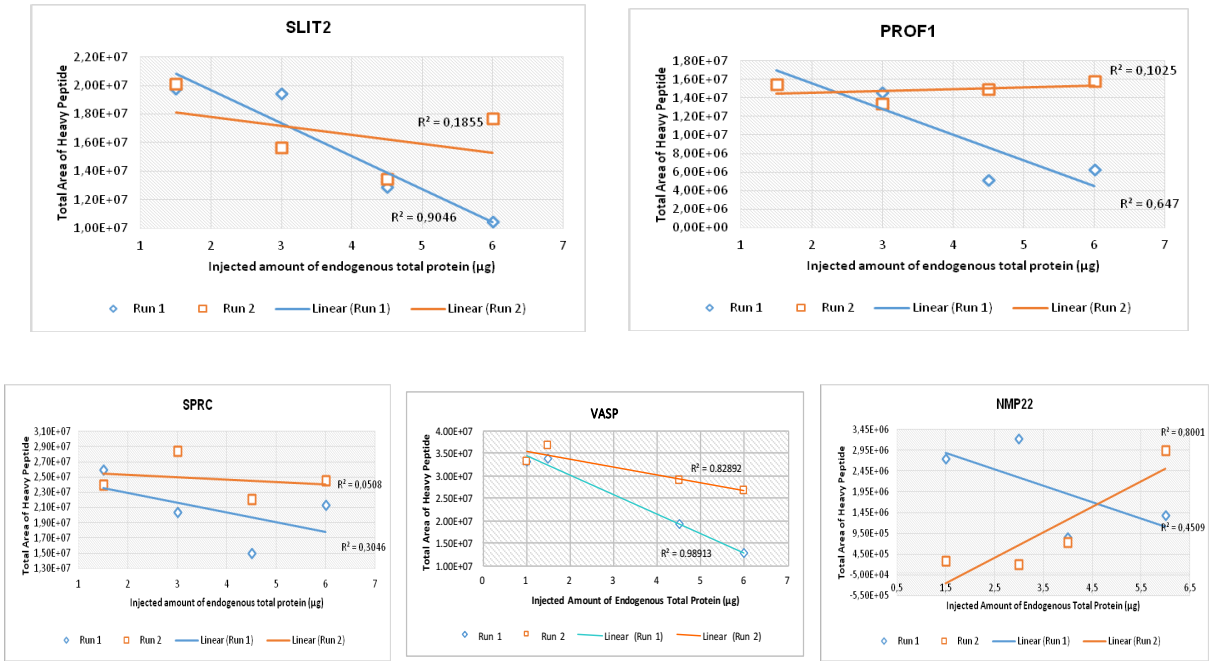


Figure 7 – Graphical representation of the heavy peptide-derived chromatographic intensity in the presence of variable amounts of endogenous urinary peptides. For the same amount heavy peptide injected (2.5 µL of a 0.002 nmol/µL stock solution in a final volume of 50 µL), variable amount of endogenous total protein, herein tested from 0.25 µg/µL (1.5 µg) to 1.0 µg/µL (6 µg) were injected.

Reproducibility

Light-to-heavy ratios obtained in the main experiment (herein named 2nd experiment in **Figure 8**) were compared to those obtained in an initial "training" one (1st experiment in **Figure 8**). Because only PROF1 and SPRC were analyzed on both experiments, comparisons were restricted to these two markers. Similarly, only samples analyzed on both experiments (matched) were used to assess assay reproducibility. The 1st assay was designed for exploratory purposes only and no data pertaining to it is herein presented, except for the data presented in this section (**3.1.6**) for reproducibility purposes. Keep in mind, however, that conditions were not perfectly matched between experiments. In particular, different chromatographic columns were used on the two experiments as the column had to be replaced between experiments, and higher amounts of heavy standard peptides with lower amounts of endogenous total protein were injected on

the 2nd one as optimization so suggested (meaning that lower light-to-heavy ratios were to be expected from the 2nd experiment). In particular, comparing the 1st to the 2nd experiment, 4 versus 3 µg total endogenous peptides were used, and 0.04 ng versus 0.0006 nmol of heavy peptide were employed (which was always higher than 0.04 ng, even for the peptide with lowest molecular-weight), respectively. Therefore, it was not surprising for the light-to-heavy peptide ratios to be always lower in the second experiment (**Figure 8**). Irrespectively, good correlation was to be expected when it comes to samples' ranking, i.e. samples displaying lower light-to-heavy ratios in the 1st experiment should have ranked lower in the 2nd one as well.

As depicted in **Figure 8**, no correlation was observed when plotting the ratio values, and neither the ratio values nor the ranks thereof (for both PROF1 and SPRC) were found to match (Pearson's correlation of 0.160 with p-value = 0.313 and Spearman's correlation of 0.107 with p-value = 0.501 for PROF1; Pearson's correlation of 0.072 with p-value = 0.649 and Spearman's correlation of -0.026 with p-value = 0.871 for SPRC). Therefore, the assay was deemed non-reproducible when dealing when urine samples. Note that, in buffer A, the reproducibility at the heavy peptides' level was not an issue (except for a shift in intensity scale which does not compromise linearity or reproducibility). While worrisome, a signal increase or decrease resulting from sample-specific interfering substances should not not compromise samples' ranking across experiments (e.g. patients with higher levels should always have higher levels). Also note that by employing a different amount of endogenous peptides in the 2nd experiment we may have unintentionally and unfortunately introduced variable amounts of unidentified interfering substances. As above mentioned, because different amounts of endogenous peptides in the 2nd experiment may have been obtained by starting with a larger urine volume for only some of the samples, the introduction of interfering substances may have been non-uniform across samples.

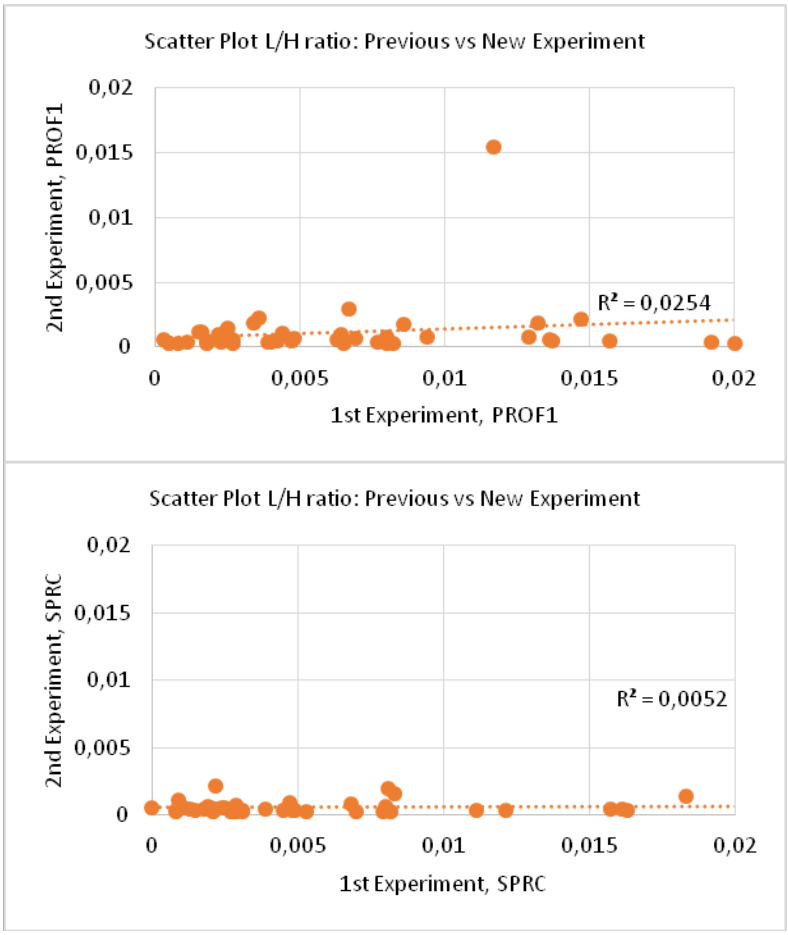


Figure 8 - Scatter plot of light-to-heavy ratio correlation for PROF1 (top) and SPRC (bottom) across two independent experiments. Different amounts of heavy and light peptides were used in the two experiments, which may account for differences at the ratio but not at the raking level.

Clinical Samples Analysis

Mass Concentration per Marker

When considering the mass concentrations of each marker across patients, most of these were noted to significantly deviate from normality using the Shapiro-Wilk test (p-value < 0.05 for all comparisons). These deviations from normality are more easily depicted graphically, specially for those groups displaying pronounced deviations

(**Figure 9**). Note that data tended to concentrate on the lower concentration range (**Figure 10**), which is not surprising as values could theoretically range from 0 to infinity (in practice up to total protein concentration). These deviations were again confirmed using Q-Q plots, but only boxplots are herein presented as these tended to be more informative. While it is not possible for concentrations to range from 0 to infinity, the trend still holds true and tells us that most patients, irrespectively of their disease group/stage, are more likely to display non-discriminating concentration values (in the lower range). Hence, the mean and the median were different within each group (mean higher than the mean) and only a minority of patients was expected to present with concentration values in the higher end of the spectrum. In **Figure 9**, the relationships between each group/marker and marker/marker combination are displayed for an initial and fast screening of any notorious patterns. Note also the absence of VASP, as it could not be discriminated from background across samples, was deemed undetectable and thus left out from the analyses.

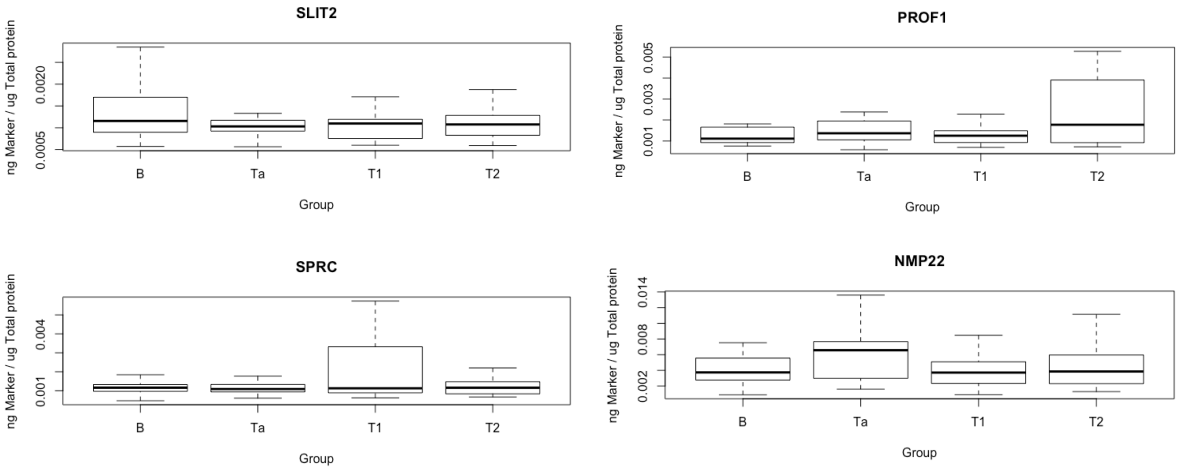


Figure 9 – Boxplots of mass concentration (ng specific marker/µg total protein) distribution for each marker across subjects (vertically) and disease group (horizontally). From left to righ in each plott: A) Benign, B) Ta, C) T1 and D) T2 groups. Note: Outliers (Q3 + 1.5*IQR, Q1 - 1.5* IQR) were automatically removed (as selected by the built-in algorithm in R). Otherwise, the scale could not fit the plots and these would be

uninterpretable. Such extreme values do not affect quantitative analyses as these were removed for this plotting purposes only but maintained for quantitative analyses. Not surprisingly, considering a positively skewed distribution, outliers fall in the region above the upper limit of Q3 + 1.5*IQR.

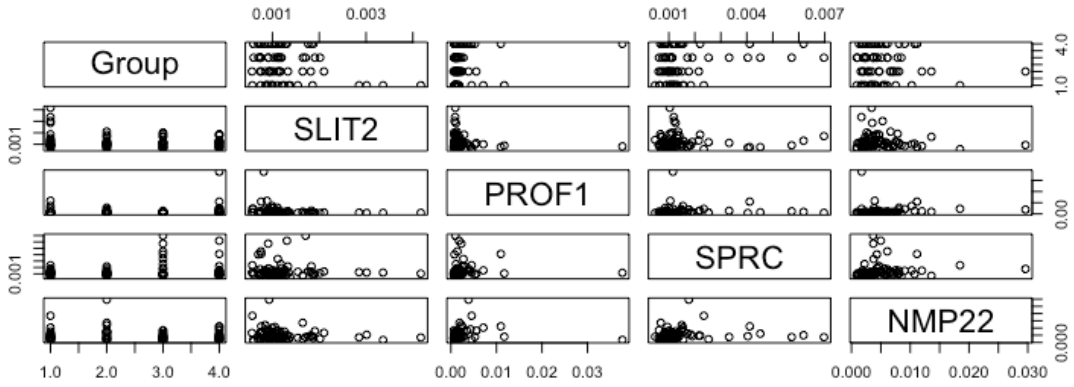


Figure 10 – General plots depicting all relationships between each disease stage and each marker and between each marker and the remaining ones. Notice that for both SPRC and NMP22, even though there are patients displaying similar values in the low concentration range, both T1 and T2 groups (herein depicted as 2.0 and 3.0, respectively) contain patients that present with considerably higher urinary concentrations of these values these 2 markers. Groups are 1.0 – B, 2.0 – T1, 3.0 – T2, 4.0 – Ta.

The median is the central value in a dataset, the cut-off for which one expects 50% of all observed values to fall above and the other 50% below it. For this reason, and because data were right-skewed, using the mean (the average value) would lead to a higher centrality value as most (>50%) patients would certainly fall below it. However, as above suggested, some but not all patients in more advanced stages of BCa did seem to display huge deviations for some of the markers under scrutiny. Hence, the median may not differ between the groups even though the higher concentration range of the spectra may point to the more advanced stages of the disease. In addition, when trying to fit a multi-marker model, the values themselves rather than their ranking are the

determining factors fro both marker selection and model performance. For these reason, both the median and the mean were compared.

Accordingly, nonparametric tests were initially performed on obtained data, using the nonparametric Mann-Whitney *U* test of the null hypothesis. This test sorted and ranked the values from any two groups to be compared and evaluated the likelihood of a randomly selected value from one sample being less or greater than a randomly selected value from another sample. With its algorithm, for each observation in one set, the number of times a value wined over any observations in the other set was counted (a value loses if the corresponding one is larger than it) and a significant Mann–Whitney *U* test is interpreted as showing a difference in the medians of the compared groups. Due to the reproducibility issues above described, only data points acquired on the same experiment were compared. As depicted in **Table 9**, no significant differences were observed when comparing the normalized (to total protein) median concentration of each marker across groups. However, such finding was not surprisingly, as exploratory analysis had suggested that if any differences existed between the groups, these would be observed only for those ~50% patients with concentration values above the median. In particular, one would expect the values in the high end of the concentration spectra to be higher for patients with malignant and advanced stages of the disease (if a marker is increased in the disease group as compared to the healthy group). For this reason, even though we were dealing with a non-normally distributed population, multiple two-tailed Student's T-tests assuming unequal variance were performed as well for each marker's abundance (**Table 10**), but making the test more stringent by adjusting the confidence level to 99%. By doing so, we were artificially prioritizing specificity over sensitivity, but if a specificity-wise robust assay could be achieved it would still be very valuable. Irrespectively of the chosen test, all samples have to be considered. Note, also, that the above remarks tell us that markers expected to be downregulated in diseased patients should not be used.

Table 6 – Mann-Whitney *U* test results

Between-Groups Comparisons
Benign vs Malignant

Marker	p-value (Mann Whitney)
[SLIT2] ng/µg Total protein	0.200491
[PROF1] ng/µg Total protein	0.540374
[SPRC] ng/µg Total protein	0.886891
[NMP22] ng/µg Total protein	0.6427739
Benign vs Ta	
Marker	p-value (Mann Whitney)
[SLIT2] ng/µg Total protein	0.3219429
[PROF1] ng/µg Total protein	0.4738751
[SPRC] ng/µg Total protein	0.6585591
[NMP22] ng/µg Total protein	0.2059581
Benign vs T1	
Marker	p-value (Mann Whitney)
[SLIT2] ng/µg Total protein	0.3219429
[PROF1] ng/µg Total protein	0.9150537
[SPRC] ng/µg Total protein	0.3526191
[NMP22] ng/µg Total protein	0.8192026
Benign vs T2	
Marker	p-value (Mann Whitney)
[SLIT2] ng/µg Total protein	0.292039
[PROF1] ng/µg Total protein	0.4071732
[SPRC] ng/µg Total protein	0.9216611
[NMP22] ng/µg Total protein	0.9216611
Ta vs T2	
Marker	p-value (Mann Whitney)
[SLIT2] ng/µg Total protein	0.9494678
[PROF1] ng/µg Total protein	0.5262435
[SPRC] ng/µg Total protein	0.8244596
[NMP22] ng/µg Total protein	0.204976

Table 7 - Unequal Variance Student's T-test (0.99 Confidence Level)

Marker	Compared Groups (x vs y)	Mean of x	Mean of y	Ratio x/y	p-value
[SLIT2] ng/µg Total protein	B vs T2	0.0015	0.0011	1.3759	0.0900
	B vs T1	0.0016	0.0011	1.3596	0.1128
	B vs Ta	0.0016	0.0011	1.3717	0.0983

[PROF1] ng/µg Total protein	B vs All Malignant	0.0015	0.0011	1.3669	0.0848
	B vs T2	0.0021	0.0044	0.4837	0.2722
	B vs T1	0.0021	0.0013	1.5970	0.2127
	B vs Ta	0.0021	0.0018	1.1955	0.6132
	B vs All Malignant	0.0021	0.0026	0.8263	0.6341
[SPRC] ng/µg Total protein	B vs T2	0.0012	0.0015	0.7698	0.2852
	B vs T1	0.0012	0.0023	0.5174	0.0334
	B vs Ta	0.0012	0.0012	1.0348	0.7878
	B vs All Malignant	0.0012	0.0017	0.7169	0.0346
[NMP22] ng/µg Total protein	B vs T2	0.0048	0.0047	1.0207	0.9307
	B vs T1	0.0048	0.0039	1.2234	0.3951
	B vs Ta	0.0048	0.0072	0.6685	0.2066
	B vs All Malignant	0.0048	0.0053	0.9146	0.6766

For the median, no significant differences were observed when comparing the normalized (to total protein) mean concentration of each marker across groups. These findings revealed by graphical representations (**Figure 9**) suggesting SLIT2 to be higher in the benign group, PROF1 higher in the T1 group and NMP22 in the T2 group were misleading. It could be the case that outliers (not represented in the figure but included in the analysis) would shift the mean of the remaining groups and level out any apparent differences. However, even if that was the case, outliers should not be removed from the analysis, as there is no analytical reason compromising their validity over the other data points.

Irrespectively, combining the four markers resulted in a better (over any of marker alone) AUC of 0.550 (±0.132), providing 70% sensitivity and 43% specificity or 85% sensitivity and 29% specificity, depending on the chosen compromise/cut-off (**Figure 11**). While far from ideal, these results still provided sensitivities higher than those achieved by other available methods such as voided urine cytology (VUC, 20-40 % [77,78]), bladder tumor associated antigen (29-91%) or NMP22 (56%) ELISAs [79,80].

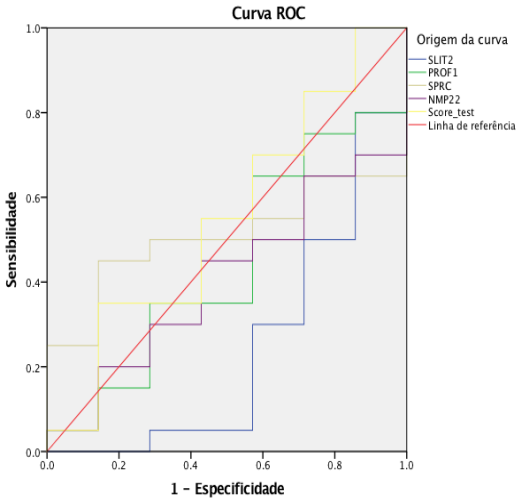


Figure 11 - ROC representation depicting the sensitivity and specificity achievable by each marker (SLIT2, PROF1, SPRC, NMP22) alone as well as combined (Score_test, yellow line) when using a mass concentration-based classifier. Reference line (red) represents the cutoff above which (the the left of which) a performance is considered superior to randomness. Parameters [cost: 0.0009765625, gamma: 1.01395947979003 , kernel: linear , epsilon: 0.1 , number of support vectors: 28.

Table 8 - Area Under the Curve Analysis

Variable Under Study	Area	Standard Error	Significance	Confidence Interval at 95%	
				Lower Limit	Upper Limit
SLIT2massConcent	0.243	0.118	0.046	0.012	0.474
PROF1massConcent	0.443	0.126	0.658	0.196	0.690
SPRCmassConcent	0.507	0.109	0.956	0.294	0.720
NMP22massConcent	0.407	0.115	0.472	0.181	0.633
Score_test	0.550	0.132	0.699	0.290	0.810

It should be recalled that the samples under study came from patients and controls with proteinuria and hematuria, factors which may compromise the ability to accurately detect the low abundance markers under scrutiny and that may influence urine protein profiles in a heterogeneous way. In fact, we had initially noted that endogenous

peptides themselves would influence the signal from the heavy peptide as well. Because the amount of light peptide is calculated based on the peak ratio to its heavy counterpart, such influence should render any quantitative measurement unreliable. Therefore, proteinuria and hematuria could be factors contributing to urine matrix effects and may compromise one's ability to trust the results. In fact, note that while for some cut-offs good sensitivity/specificity could be achieved, the AUC was still very low. Such observation told us that most possible cut-off values were non-discriminatory and that the apparently discriminatory ones were most likely a result of randomness.

At this point, it was evident that the assay was not robust enough to overcome the limitations resulting from the limitating inherent urinary matrix effects, which were sample-specific and not predictable. Even though one could think of further purification as a possible means for attenuating matrix effects, the results of adding other steps are also unpredictable and may not be advantageous. For instance, it is known that protein depletion strategies seem to provide no added detection value over unfractionated urine samples [81]. Also, if these were to be required for the herein presented purpose (accurate targeted quantification), they would make the assay impracticable in the clinical setting. Not only that, one of the major drawbacks of high-abundance protein depletion is the loss of low-abundance species, as the former ones are known to bound and carry the latter ones [82]. For this reason, for the time being, we ought not to employ such strategies when aiming at targeted protein quantification, as we do not know to what extent this practice would affect different samples. Plus, we did not know which species to deplete, as one could not be sure if the matrix effects were due to urinary proteins or some other non-protein species.

While mass concentration of specific biomarker per total amount of protein (i.e. total amount of biomarker, as the same amount of total protein was injected for each sample) should be deemed a more reliable way to quantify chromatographic peaks, it relies on the light-to-heavy ratios. It is the case, however, that the same retention time window over which the signal for the heavy peptide was measured could not be used across samples when trying to maximize the match between light and heavy peptides. For instance, wider retention time windows had to be employed for both heavy and light peptides in more spread samples, which artificial increases the light-to-heavy ratio for

samples with more spread light chromatograms. Therefore, this could mean that differences at the light peptide level were hampered by using different analysis criteria for the heavy peptides across samples when aiming at light/heavy match maximization. We thus opted to measure and compare the total area for the light peptide peaks only, possibly avoiding the inherent bias of ratio calculations when the denominator is unreliable.

Clinical Samples Analysis: Total Chromatogram Area

The total chromatographic area of each endogenous peptide (**Figure 12**) was compared between all groups, and a SVM method was employed in order to create a model combining all four markers (normalized/total area refers to the sum of all transitions for that particular precursor). Initially, by looking at the distribution of the light peptide area values across groups (**Figure 12**), no immediate differences could be perceived. Such lack of remarkable differences was again confirmed by carefully inspecting the raw values (**Table 11**). Note, also, that the heavy peptide signal should be stable across all groups. However, while not so pronounced as the ones observed during the QCs, considerable differences were still noted for the signal intensities derived from the heavy peptides as well. Also, major deviations at the level of the heavy peptide-derived chromatogram total areas took place across patients for all groups. Such variability is more easily perceived by looking at the large standard deviations inside each group in **Figure 13**, and this observation reinforced the above presented evidences suggesting urine matrix to compromise the ability of MRM assays to achieve accurate quantification.

These shifts may reflect and be the result of different protein profiles if one considers the possibility of quenching, summing or crosstalk effects. Also, peptides are assumed to be fully reduced and to be free from interactions with other peptides, but this assumption may not be absolutely true. We don't know if the presence of urinary metabolites not fully removed during sample preparation might have contributed to the highly variable interaction/influence that can be observed across samples. In addition, the issue is certainly made more complicated to interpret, as both the disease and

control groups consisted of a very heterogeneous population including individuals under diverse drug regimens, with several distinct comorbidities (e.g. prostate cancer, urinary track infections) and with variable degrees of hematuria. Together, these factors may have contributed to variability at the level of urinary molecules other than the peptide markers under investigation. Accordingly, recall that when varying the amount (concentration) of total endogenous peptides derived from the same sample, this phenomenon was observed, suggesting that variable amounts of interfering species was the most likely predominating factor under question.

Salt concentrations were most likely not an issue because TCA precipitation followed by washing steps must have removed most salts, as these are very well soluble in aqueous solutions. In addition, the pH as been previously shown the have no significant effects [83], even more so when one considers the precipitation and desalting steps herein performed. Interestingly, urine specific gravity is known to influence the percentage recovery of ELISAs to a significant extend. Unfortunately, its influence is still somewhat unpredictable and not amenable to be accounted for [83]. Because several substances dictate urine specific gravity (urea, sodium, potassium, and ammonium salts), it is probable for variable composition of these substances across subjects to have accounted for such unpredictability.

Urine is also rich in lipophilic substances, which may remain after TCA precipitation. Still, TCA was chosen as the precipitant because it is known to provide higher (protein) yields and its typical drawback in proteomics, protein denaturation, was not an issue for our analyses. One may think of organic solvent (e.g. acetone, ethanol) precipitation as a plausible alternative to try to overcome the herein presented challenges. However, with these, protein recovery yields would be more variable, it would increase the turnaround time in the clinical setting and would require additional time with the mass spectrometer for our group to compare different methods, which was limited during this project. Still, frequently observed tailings in the chromatograms do support the hypothesis of lipophilic interfering substances, but even if an efficacious alternative was immediately available it should be carefully pondered, as any extra step might also introduce artifacts.

At this point, it became clear that any comparisons between patients or groups thereof would be faulty and always biased in an unpredictable way. Take, for instance, the example of SLIT2. Even though an initial screening suggested its light/endogenous-derived signal to increase more than 2-fold from the B to T2 group, the signal from the internal heavy peptide control more than doubled as well. Consequently, even though SLIT2 appeared to be increasing when looking at its endogenous signal only, the ratio to its heavy counterpart (which should follow the same trend as the exact same amount of heavy peptide was injected in all samples) actually decreased (**Figure 15**). Similarly, the amount of PROF1 seemed to remain constant when comparing B and T2 groups, the extremes of the disease spectrum. In turn, when observing the light-to-heavy ratio it was suggested that PROF1 actually increased 50-fold from the benign to the T2 group, a discrepancy that could not be taken lightly. Note, however, that in the case of PROF1, the signal from the heavy peptide remained (apparently) unaffected. How may this discrepancy be then explained? It would be possible if those samples with the highest values ("values", not necessarily "amounts of") of PROF1 were the ones displaying lower levels of heavy peptide-derived intensities, so that major outliers would result in the higher end of the concentration ratios' spectrum, significantly affecting the ratio of the means.

Still, when 4 (SLIT2, SPRC, PROF1 and NMP22) markers together were combined on a total chromatographic area-based predictive model, an AUC of 0.814 (±0.083), p-value = 0.015, with a sensitivity of ~70% and specificity of ~100% or with a sensitivity of ~80% and specificity of ~57% was achieved, depending on the chosen compromise (**Figure 16**). Even if one considered the classifier to be performing sub-optimally (due to variability in heavy peptide measurements), these values still suggested the combined markers in cause to be (apparently) suitable for the intended purpose and the assay to be robust enough to detect differences across the compared groups. Moreover, it should be noted that both the training and test sets used for training and testing the model (in both marker concentration and area-based models) were smaller than required to achieve a model with significant predictive capacities. On the other hand, as the small data set had a 2-fold uncertainty with it associated, it also meant that any apparent patterns may have appeared by chance alone. Accordingly, adding the poor statistical power together

with the MS analytical challenges observed when dealing with urine matrix compromised our confidence in the observed results.

Supposing the observed performance was in fact achievable (from a theoretical point-of-vie), BCa would be significantly better addressed if detected early, allowing 5-year survival rates as high as 94%. As a result, high sensitivity tests are paramount, and while this urgent need remains to be fulfilled, we have shown that MRM assays may (in theory) achieve the required sensitivity. Also, frequent long-term surveillance and repeated evaluation are needed to prevent disease progression, but relatively cheap and cost-effective tests that can be repeatedly and affordably performed are lacking. Again reinforcing the advantages of MRM assays over conventional testing, the former are cheaper than the latter because in the long run implementation costs end-up being returned. If one also takes into account the invasiveness of and complications associated with current diagnostic procedures for BCa, it becomes evident that a urinary protein MRM-based classifier would be a perfect test.

The gold standard for BCa diagnosis in the clinics remains cystoscopy examination of the bladder coupled with voided urine cytology (cytological examination of cellular material in urine) [77,84]. Cystoscopy is uncomfortable, costly and invasive, requiring anesthetization of the patient and being associated with major complications. While it achieves optimal sensitivity and specificity, it cannot be repeatedly performed for obvious reasons. In turn, VUC relies on the microscopic visualization of shed cancer cells into voided urine, but it fails to detect the disease in its early stages and achieves only modest sensitivities of 20-40% [77,78]. Also, in contrast with SRM/MRM assays, it is very subjective (observer-dependent) [77,84]. While VUC can achieve high specificity (~80%), this parameter is not an issue as there are other non-invasive assays capable of achieving great specificity. Therefore, these could theoretically be easily coupled to a sensitive MRM assay. Still, such MRM assay would most likely provide comparable specificity if the correct markers are chosen.

While there are some commercial ELISA-based tests capable of mirroring the standard tests' performance, such tests are not robust enough to overcome the complexity of urine matrix. For instance, bladder tumor associated antigen can achieve sensitivities from 29% to 91% and specificities from 56% to 86%, but the results are highly influenced

by co-morbidities. As a result, factors such as hematuria make the assays unreliable [79]. Similarly, even though ELISA-based NMP22 screening achieves maximal sensitivity of 56% and maximal specificity of 87% (not simultaneously, though), its performance is compromised by several comorbidities [80]. Therefore, protein MS-based assays preformed post-protein isolation are expected to be less affected by the multiple interfering components of urine samples. However, as we have observed, even for MS SRM/MRM-based assays this may not be as straightforward as one would expect.

It is to be acknowledged that the end result of the MRM assay herein developed was far from perfect, as major issues were unveiled regarding its accuracy in urine matrix. Still, these were better than the ones provided by the currently available tests. At this point, it is not possible to pinpoint with confidence what caused the observed drifting in the heavy peptides-derived signals (and an accompanying bias in the signal from the light counterparts if one chooses to compare the ratios). Nevertheless, results are somewhat encouraging as these outperformed those previously attained using ELISAs and higher recoveries were still achieved.

However, as it should by now be evident, one cannot fully trust these findings due to the lack of heavy peptide signal robustness in urine matrixes. In fact, as pointed out by Marc Twain, "There are three kinds of lies: lies, damned lies and statistics" (*Mark Twain's Own Autobiography: The Chapters from the North American Review*). Accordingly, one cannot be certain of what is apparently being measured during these or other MRM assays performed on urinary proteins unless one assures the absence of urinary matrix interferences. Therefore, the use of response measurements on interference-free MRM transitions for protein quantitation when dealing with urine samples is still not feasible, as the matrix itself compromises even the most accurate and reliable available system.

We have attributed the main limitation of the developed assay, namely lack of accuracy and reproducibility of the heavy standards in urine samples, to the so-called “matrix effects”. But what does this mean and what could one do to address this issue? Matrix effects consist of the effect that all substances (e.g. salts, metabolites, carbohydrates, proteins, xenobiotics) other than the analyte of interest in a specific sample cause on an analytical assay [85]. For instance, lipids and other lipophilic plasma constituents may quench detectors and hamper the detection of other substances by altering its

hydrophobicity. However, depleting all interfering substances would not be possible. It would perhaps be more promising to enrich for targeted proteins/peptides, but methods for such purpose are lacking. Also, MRM may suffer from a "cross-talk" phenomenon, whereby related substances with similar retention times can "cross-talk" with the internal standards and interfere with its signal [86], which most certainly took place during this assay development as evidenced by the tailings on the extracted ion chromatograms.

Typically, matrix effects are not considered for method validation purposes as long as these do not influence assay reproducibility and linearity, but while instrumentation may still be guaranteed to be operating in an acceptable way (according to the apparatus specifications), suitability for the intended purpose can not be assured. Accordingly, a validated method may be operating optimally, but matrix effects may still compromise sensitivity even without compromising assay reproducibility and linearity. As a result, if biological samples to analyze by LC-MS/MS are not pure enough, elevated background and signal suppression/enhancement can occur.

In order to attenuate matrix effects and increase the signal-to-noise ratio, chromatographic purification should be optimized on a case-by-base basis, as each biological sample presents with particular challenges. One way of testing if sensitivity is being lost during chromatographic resolution would comprehend the comparison of the signal-to-noise ratio when the sample is injected first into the chromatographic column versus direct injection into the analyzer source [87]. However, for complex samples such as urine this would not be feasible (these cannot be directly injected into the analyzer source). Alternatively, one could analyze and compare the signals derived from the analyte spiked into blank sample versus the signal derived from analyte spiked into buffer. However, this method requires the concentration of analyte in the sample to be known a priori as well as both pure analyte and "blank" sample to be available [88]. As a last alternative and as herein performed, matrix effects can be assessed (but not accounted for) by comparing the signal derived from an isotope labeled internal standard spiked into the sample versus the signal resulting from spiking it into solvent. By doing so, we have not only observed the signals to be considerable different but also the linearity and reproducibility to be lost.

In what may seem an obvious choice, adding another dimension to chromatographic separations (in-line) or another decoupled liquid or solid phase purification step (off-line) should improve chromatographic resolution and improve downstream ionization efficiency (e.g. organic extraction to remove red cell-derived substance from hematuria patients). However, as previously mentioned, these additional steps would increase the cost, the turnaround time and the complexity of sample preparation, which may compromise its introduction into the clinical setting. Alternatively, diluting urine samples has been shown to attenuate the coefficient of variation (most likely by diminishing the concentration of interfering substances) and improved the quantitative recovery (up to two-fold). However, one should always consider the the lower LOD will also increase, as the amount of analyte is decreased [83].

The presence of late eluting peaks may also compromise peptide quantification, as these can appear in and interfere with subsequent samples up for a long time (e.g. up to 40 minutes), thus compromising the analysis of several samples [85]. However, increasing the organic solvent phase between runs has been shown to ameliorate this issue. Irrespectively, assuming that these limitations can be overcome and sensitivity, linearity and reproducibility can be improved by attenuating matrix effects, one should still ask if such enhancements result from improved chromatographic separations or optimized ionization efficiency. One way to do this, would be by comparing the signal derived from an isotope labeled internal standard spiked into the sample versus the signal resulting from spiking it into solvent (as above described). This time, comparisons should take place at both the product and precursor levels [88]. After accounting for improved chromatographic recovery, if these precursor and product ions both showed improvements to a similar extend, one may safely assume such improvements to result from optimized ionization efficiency. However, if the improvement at the product ion level is superior to the one observed at the level of its precursor, improved fragmentation (or shifted towards other fragmentation pathways, as interfering substance may also influence these) has also to be acknowledged.

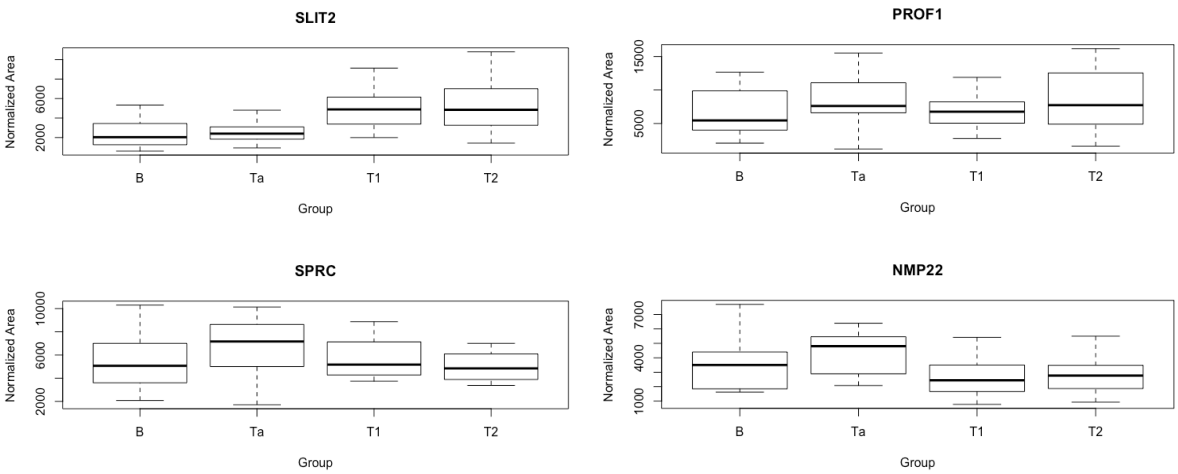
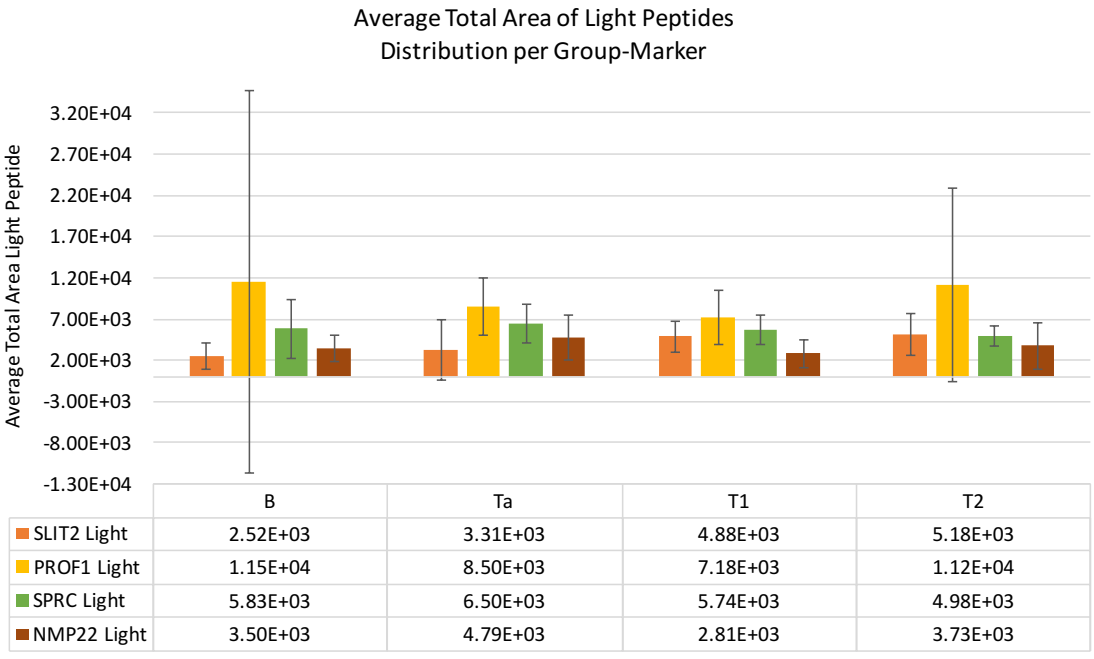


Figure 12 - Boxplots of normalized endogenous peptide-derived chromatogram area distribution for each marker across subjects (vertically) and disease group (horizontally). From left to right in each plot: A) Benign, B) Ta, C) T1 and D) T2 groups. Note: Outliers (Q3 + 1.5*IQR, Q1 - 1.5* IQR) were automatically removed (as selected by the built-in algorithm in R). Such extreme values do not affect quantitative analyses as these were removed for this plotting purposes only but maintained for quantitative analyses.

Table 8 - Total Area Results for all Groups and all Peptides

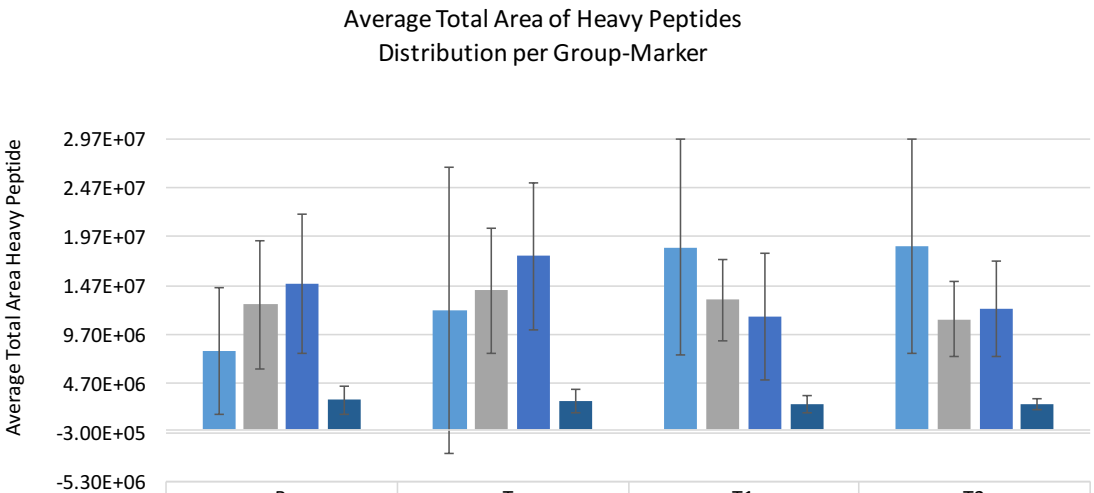
	SLIT2					
	Light Area		Heavy Area		Ratio	
	Average	SD	Average	SD	Average	SD
B	2,52E+03	1,61E+03	8,02E+06	6,47E+06	4,18E-04	2,70E-04
Ta	3,31E+03	3,67E+03	1,21E+07	1,46E+07	3,05E-04	1,08E-04
T1	4,88E+03	1,91E+03	1,86E+07	1,10E+07	3,09E-04	1,17E-04
T2	5,18E+03	2,50E+03	1,88E+07	1,09E+07	3,04E-04	1,02E-04
	PROF1					
	Light Area		Heavy Area		Ratio	
	Average	SD	Average	SD	Average	SD
B	1,15E+04	2,31E+04	1,27E+07	6,57E+06	8,75E-04	1,11E-03
Ta	8,50E+03	3,53E+03	1,42E+07	6,39E+06	7,32E-04	5,15E-04
T1	7,18E+03	3,30E+03	1,33E+07	4,15E+06	5,47E-04	2,30E-04
T2	1,12E+04	1,18E+04	1,13E+07	3,79E+06	1,81E-03	3,45E-03
	SPRC					
	Light Area		Heavy Area		Ratio	

	Average	SD	Average	SD	Average	SD
B	5,83E+03	3,57E+03	1,49E+07	7,08E+06	4,08E-04	1,46E-04
Ta	6,50E+03	2,33E+03	1,77E+07	7,46E+06	3,95E-04	1,35E-04
T1	5,74E+03	1,78E+03	1,16E+07	6,48E+06	7,89E-04	6,67E-04
T2	4,98E+03	1,24E+03	1,23E+07	4,84E+06	5,31E-04	4,70E-04
	NMP22					
	Light Area		Heavy Area		Ratio	
	Average	SD	Average	SD	Average	SD
B	3,50E+03	1,58E+03	3,05E+06	1,43E+06	1,46E-03	1,18E-03
Ta	4,79E+03	2,67E+03	2,90E+06	1,24E+06	2,18E-03	2,03E-03
T1	2,81E+03	1,63E+03	2,54E+06	8,96E+05	1,19E-03	6,70E-04
T2	3,73E+03	2,82E+03	2,60E+06	5,93E+05	1,43E-03	9,25E-04



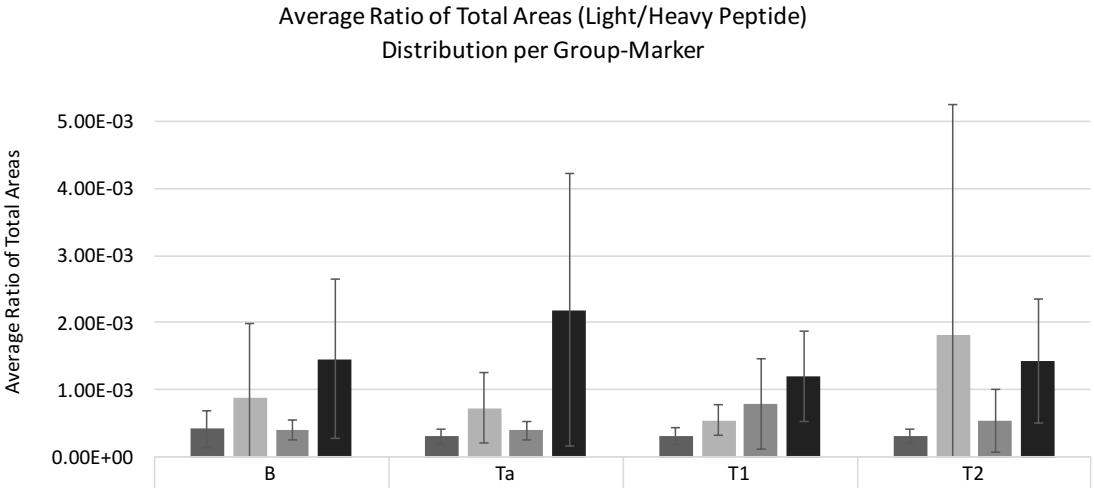
	B	Ta	T1	T2
SLIT2 Light	2.52E+03	3.31E+03	4.88E+03	5.18E+03
PROF1 Light	1.15E+04	8.50E+03	7.18E+03	1.12E+04
SPRC Light	5.83E+03	6.50E+03	5.74E+03	4.98E+03
NMP22 Light	3.50E+03	4.79E+03	2.81E+03	3.73E+03

Figure 13 – Graphical representation of average (± St. Dev.) total chromatographic area for SLIT2, PROF1, SPRC and NMP22 endogenous peptides across BCa disease groups.



	B	Ta	T1	T2
SLIT2 Heavy	8.02E+06	1.21E+07	1.86E+07	1.88E+07
PROF1 Heavy	1.27E+07	1.42E+07	1.33E+07	1.13E+07
SPRC Heavy	1.49E+07	1.77E+07	1.16E+07	1.23E+07
NMP22 Heavy	3.05E+06	2.90E+06	2.54E+06	2.60E+06

Figure 14 - Graphical representation of average (± St. Dev.) total chromatographic area for SLIT2, PROF1, SPRC and NMP22 heavy standard control peptides across BCa disease groups.



	B	Ta	T1	T2
SLIT2 Ratio	4.18E-04	3.05E-04	3.09E-04	3.04E-04
PROF1 Ratio	8.75E-04	7.32E-04	5.47E-04	1.81E-03
SPRC Ratio	4.08E-04	3.95E-04	7.89E-04	5.31E-04
NMP22 Ratio	1.46E-03	2.18E-03	1.19E-03	1.43E-03

Figure 15 - Graphical representation of average (± St. Dev.) light-to-heavy ratios of chromatographic areas for SLIT2, PROF1, SPRC and NMP22 across BCa disease groups.

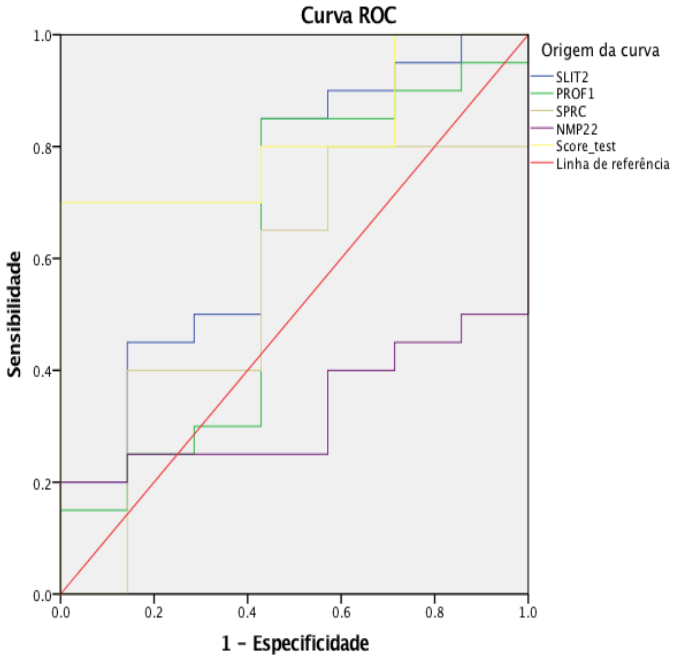


Figure 16 – ROC representation depicting the sensitivity and specificity achievable by each marker (SLIT2, PROF1, SPRC, NMP22) alone as well as combined (Score_test, yellow line) when using a total chromatographic area-based classifier. Reference line (red) represents the cutoff above which (the the left of which) a performance is considered superior to randomness. Parameters [cost: 0.0009765625 , gamma: 1.01395947979003 , kernel: sigmoid , epsilon: 0.1 , number of support vectors: 27].

Conclusions

At the individual biomarker level, the MRM assay herein developed for urine profiling provided comparable-to-superior analytical performance (accuracy or percentage recovery) as compared to ELISA assays. The combination of SLIT2, PROF1, SPRC and NMP22 in a 4-marker classifier also resulted in comparable-to-superior clinical performance (~70% sensitivity with ~100% specificity or ~80% sensitivity with ~57% specificity, depending on the chosen cut-off value) over previously developed assays. Notably, the resulting classifier provided sensitivities higher than those achieved by

other available methods such as voided urine cytology, bladder tumor associated antigen or NMP22 ELISAs, therefore addressing the main limitation thereof. However, interference-free measurements still could not be guaranteed. In particular, the heavy peptide-derived chromatograms were shifted in a highly variable way across samples, which suggests sample-specific factors to bias the analytical performance of the MRM assay developed.

Future Directions

Larger cohorts will ideally be tested in order to surpass statistical limitations, as only 73 samples were available. The molecular mediators of urinary matrix effects will have to be identified if urine samples are to be exploited for biomarker profiling using MRM assays. Afterwards, a protocol for their efficient removal will have to be developed in order to overcome such matrix effects. New approaches have been recently presented (e.g. see P Bastos *et al*. 2017 [89]) and more are expected to emerge. In addition, turnaround time and economical costs will have to be re-evaluated if additional sample handling steps turn out to be required for MRM introduction into the clinical setting. Some BCa-specific urinary markers could be unveiled by meta-analysis (see Bastos *et al. Under Review* on section II), namely Peroxisomal acyl-coenzyme A oxidase 1, Aspartate aminotransferase mitochondrial, Aldo-keto reductase family 1 member C2, Adipocyte plasma membrane-associated protein, Adipocyte enhancer-binding protein 1, C1q subunits and C4b-binding protein. These are other alternative promising markers for validation may also be exploited and are expected to contribute with increased specificity for future multimarker panels.

REFERENCES

1. Shi T, Su D, Liu T, Tang K, Camp DG, Qian WJ, et al. Advancing the sensitivity of selected reaction monitoring-based targeted quantitative proteomics. Proteomics. 2012. p. 1074–92.
2. Anderson L, Hunter CL. Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. Mol. Cell. Proteomics. 2006;5:573–88.
3. Picotti P, Bodenmiller B, Mueller LN, Domon B, Aebersold R. Full Dynamic Range Proteome Analysis of S. cerevisiae by Targeted Proteomics. Cell. 2009;138:795–806.
4. Addona TA, Abbatiello SE, Schilling B, Skates SJ, Mani DR, Bunk DM, et al. Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. Nat. Biotechnol. 2009;27:633–41.
5. Beasley-Green A, Burris NM, Bunk DM, Phinney KW. Multiplexed LC–MS/MS Assay for Urine Albumin. J. Proteome Res. American Chemical Society; 2014;13:3930–9.
6. Chen CL, Lai YF, Tang P, Chien KY, Yu JS, Tsai CH, et al. Comparative and targeted proteomic analyses of urinary microparticles from bladder cancer and hernia patients. J. Proteome Res. 2012;11:5611–29.
7. Chen YT, Chen HW, Domanski D, Smith DS, Liang KH, Wu CC, et al. Multiplexed quantification of 63 proteins in human urine by multiple reaction monitoring-based mass spectrometry for discovery of potential bladder cancer biomarkers. J. Proteomics. 2012;75:3529–45.
8. Zhang H, Cao J, Li L, Liu Y, Zhao H, Li N, et al. Identification of urine protein biomarkers with the potential for early detection of lung cancer. Sci. Rep. 2015;5:11805.
9. Simon R, Lemoine J, Fonbonne C, Jaffuel A, Léonard J-F, Gautier J-C, et al. Absolute quantification of podocin, a potential biomarker of glomerular injury in human urine, by Liquid Chromatography - Multiple Reaction Monitoring cubed Mass spectrometry. J. Pharm. Biomed. Anal. 2014;
10. Mermelekas G, Vlahou A, Zoidakis J. SRM/MRM targeted proteomics as a tool for biomarker validation and absolute quantification in human urine. Expert Rev. Mol. Diagn. 2015;15:1441–54.
11. Deutsch EW, Lam H, Aebersold R, Anderson L, Hunter C, Craig R, et al. PeptideAtlas: a resource for target selection for emerging targeted proteomics workflows. EMBO Rep. 2008;9:429–34.
12. Craig R, Cortens JP, Beavis RC. Open source system for analyzing, validating, and storing protein identification data. J. Proteome Res. 2004;3:1234–42.
13. Shadforth I, Xu W, Crowther D, Bessant C. GAPP: A fully automated software for the confident identification of human peptides from tandem mass spectra. J. Proteome Res. 2006;5:2849–52.

14. Kuster B, Schirle M, Mallick P, Aebersold R. Scoring proteomes with proteotypic peptide probes. Nat. Rev. Mol. Cell Biol. 2005;6:577–83.
15. Mallick P, Schirle M, Chen SS, Flory MR, Lee H, Martin D, et al. Computational prediction of proteotypic peptides for quantitative proteomics. Nat. Biotechnol. 2007;25:125–31.
16. Fusaro V a, Mani DR, Mesirov JP, Carr S a. Prediction of high-responding peptides for targeted protein assays by mass spectrometry. Nat. Biotechnol. 2009;27:190–8.
17. Picotti P, Aebersold R, Domon B. The implications of proteolytic background for shotgun proteomics. Mol. Cell. Proteomics. 2007;6:1589–98.
18. Brose K, Bland KS, Wang KH, Arnott D, Henzel W, Goodman CS, et al. Slit proteins bind Robo receptors and have an evolutionarily conserved role in repulsive axon guidance. Cell. 1999;96:795–806.
19. Huang T, Kang W, Cheng ASL, Yu J, To KF. The emerging role of Slit-Robo pathway in gastric and other gastro intestinal cancers. BMC Cancer. 2015;15:950.
20. Stein E, Tessier-Lavigne M. Hierarchical organization of guidance receptors: silencing of netrin attraction by slit through a Robo/DCC receptor complex. Science. 2001;291:1928–38.
21. Castets M, Broutier L, Molin Y, Brevet M, Chazot G, Gadot N, et al. DCC constrains tumour progression via its dependence receptor activity. Nature. 2011;482:534–7.
22. Prasad A, Fernandis AZ, Rao Y, Ganju RK. Slit Protein-mediated Inhibition of CXCR4-induced Chemotactic and Chemoinvasive Signaling Pathways in Breast Cancer Cells. J. Biol. Chem. 2004;279:9115–24.
23. Stella MC, Trusolino L, Comoglio PM. The Slit/Robo system suppresses hepatocyte growth factor-dependent invasion and morphogenesis. Mol. Biol. Cell. 2009;20:642–57.
24. Dallol A, Da Silva NF, Viacava P, Minna JD, Bieche I, Maher ER, et al. SLIT2, a human homologue of the Drosophila Slit2 gene, has tumor suppressor activity and is frequently inactivated in lung and breast cancers. Cancer Res. 2002;62:5874–80.
25. Xu Y, Li WL, Fu L, Gu F, Ma YJ. Slit2/Robo1 signaling in glioma migration and invasion. Neurosci. Bull. 2010. p. 474–8.
26. Leek RD, Landers RJ, Harris AL, Lewis CE. Necrosis correlates with high vascular density and focal macrophage infiltration in invasive carcinoma of the breast. Br. J. Cancer. 1999;79:991–5.
27. Zhu J, Jiang Z, Gao F, Hu X, Zhou L, Chen J, et al. A systematic analysis on DNA methylation and the expression of both mRNA and microRNA in bladder cancer. PLoS One. 2011;6.
28. Roperch J-P, Grandchamp B, Desgrandchamps F, Mongiat-Artus P, Ravery V, Ouzaid I, et al. Promoter hypermethylation of HS3ST2, SEPTIN9 and SLIT2 combined with FGFR3 mutations as a sensitive/specific urinary assay for diagnosis and surveillance in patients with low or high-risk non-muscle-invasive bladder cancer. BMC Cancer. 2016;16:704.
29. Zoidakis J, Makridakis M, Zerefos PG, Bitsika V, Esteban S, Frantzi M, et al. Profilin 1

is a potential biomarker for bladder cancer aggressiveness. Mol. Cell. Proteomics. 2012;11:M111.009449.
30. Schlake A, Crispen PL, Cap AP, Atkinson T, Davenport D, Preston DM. NMP-22, urinary cytology, and cystoscopy: A 1 year comparison study. Can. J. Urol. 2012;19:6345–50.
31. Wirth MP, Hakenberg OW. Detection of bladder cancer using a point-of-care proteomic assay. Eur. Urol. 2006. p. 196.
32. Grossman HB, Soloway M, Messing E, Katz G, Stein B, Kassabian V, et al. Surveillance for recurrent bladder cancer using a point-of-care proteomic assay. JAMA. 2006;295:299–305.
33. Tritschler S, Scharf S, Karl A, Tilki D, Knuechel R, Hartmann A, et al. Validation of the Diagnostic Value of NMP22 BladderChek Test as a Marker for Bladder Cancer by Photodynamic Diagnosis. Eur. Urol. 2007;51:403–8.
34. Shariat SF, Marberger MJ, Lotan Y, Sanchez-Carbayo M, Zippe C, Lüdecke G, et al. Variability in the Performance of Nuclear Matrix Protein 22 for the Detection of Bladder Cancer. J. Urol. 2006;176:919–26.
35. Said N, Frierson HF, Sanchez-Carbayo M, Brekken RA, Theodorescu D. Loss of SPARC in bladder cancer enhances carcinogenesis and progression. J. Clin. Invest. 2013;123:751–66.
36. Yamanaka M, Kanda K, Li NC, Fukumori T, Oka N, Kanayama HO, et al. Analysis of the gene expression of SPARC and its prognostic value for bladder cancer. J Urol. 2001;166:2495–9.
37. Latosinska A, Makridakis M, Frantzi M, Borràs DM, Janssen B, Mullen W, et al. Integrative analysis of extracellular and intracellular bladder cancer cell line proteome with transcriptome: improving coverage and validity of –omics findings. Sci. Rep. 2016;6:25619.
38. Pisitkun T, Shen R-F, Knepper M a. Identification and proteomic profiling of exosomes in human urine. Proc. Natl. Acad. Sci. U. S. A. 2004;101:13368–73.
39. Welton JL, Khanna S, Giles PJ, Brennan P, Brewis I a, Staffurth J, et al. Proteomics analysis of bladder cancer exosomes. Mol. Cell. Proteomics. 2010;9:1324–38.
40. Théry C. Exosomes: secreted vesicles and intercellular communications. Biol. Reports. 2011;33410:15–3.
41. Frantzi M, Bhat A, Latosinska A. Clinical proteomic biomarkers: relevant issues on study design & technical considerations in biomarker development. Clin. Transl. Med. 2014;3:7.
42. Selevsek N, Matondo M, Carbayo MS, Aebersold R, Domon B. Systematic quantification of peptides/proteins in urine using selected reaction monitoring. Proteomics. 2011;11:1135–47.
43. Olszowy P, Buszewski B. Urine Sample Preparation for Proteomic Analysis. J. Sep. Sci. 2014;1–26.

44. De Marchi T, Kuhn E, Carr SA, Umar A. Antibody-Based Capture of Target Peptides in Multiple Reaction Monitoring Experiments. Methods Mol Biol. 2015. p. 123–35.
45. Picotti P, Rinner O, Stallmach R, Dautel F, Farrah T, Domon B, et al. High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. Nat. Methods. 2010;7:43–6.
46. Chernushevich I V., Loboda A V., Thomson BA. An introduction to quadrupole-time-of-flight mass spectrometry. J. Mass Spectrom. 2001;36:849–65.
47. Domanski D, Percy AJ, Yang J, Chambers AG, Hill JS, Freue GVC, et al. MRM-based multiplexed quantitation of 67 putative cardiovascular disease biomarkers in human plasma. Proteomics. 2012;12:1222–43.
48. Ramirez-Correa GA, Martinez-Ferrando MI, Zhang P, Murphy AM. Targeted proteomics of myofilament phosphorylation and other protein posttranslational modifications. Proteomics - Clin. Appl. 2014. p. 543–53.
49. Roepstorff P, Fohlman J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. Biomed. Mass Spectrom. 1984;11:601.
50. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc. Natl. Acad. Sci. U. S. A. 2003;100:6940–5.
51. Dupuis A, Hennekinne JA, Garin J, Brun V. Protein Standard Absolute Quantification (PSAQ) for improved investigation of staphylococcal food poisoning outbreaks. Proteomics. 2008;8:4633–6.
52. Zhang H, Liu Q, Zimmerman LJ, Ham A-JL, Slebos RJC, Rahman J, et al. Methods for peptide and protein quantitation by liquid chromatography-multiple reaction monitoring mass spectrometry. Mol. Cell. Proteomics. 2011;10:M110.006593.
53. Picotti P, Aebersold R. Selected reaction monitoring–based proteomics: workflows, potential, pitfalls and future directions. Nat. Methods. 2012;9:555–66.
54. Keshishian H, Addona T, Burgess M, Kuhn E, Carr SA. Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. Mol. Cell. Proteomics. 2007;6:2212–29.
55. Brownridge P, Holman SW, Gaskell SJ, Grant CM, Harman VM, Hubbard SJ, et al. Global absolute quantification of a proteome: Challenges in the deployment of a QconCAT strategy. Proteomics. 2011. p. 2957–70.
56. Beynon RJ, Doherty MK, Pratt JM, Gaskell SJ. Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. Nat. Methods. 2005;2:587–9.
57. Holzmann J, Pichler P, Madalinski M, Kurzbauer R, Mechtler K. Stoichiometry determination of the MP1-p14 complex using a novel and cost-efficient method to produce an equimolar mixture of standard peptides. Anal. Chem. 2009;81:10254–61.
58. Kuzyk M a, Smith D, Yang J, Cross TJ, Jackson AM, Hardie DB, et al. Multiple reaction monitoring-based, multiplexed, absolute quantitation of 45 proteins in human plasma.

Mol. Cell. Proteomics. 2009;8:1860–77.
59. MacLean B, Tomazela DM, Abbatiello SE, Zhang S, Whiteaker JR, Paulovich AG, et al. Effect of collision energy optimization on the measurement of peptides by selected reaction monitoring (SRM) mass spectrometry. Anal. Chem. 2010;82:10116–24.
60. Stahl-Zeng J, Lange V, Ossola R, Eckhardt K, Krek W, Aebersold R, et al. High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. Mol. Cell. Proteomics. 2007;6:1809–17.
61. Kiyonami R, Schoen A, Prakash A, Peterman S, Zabrouskov V, Picotti P, et al. Increased selectivity, analytical precision, and throughput in targeted proteomics. Mol. Cell. Proteomics. 2010;in press.
62. Brusniak M-YK, Kwok S-T, Christiansen M, Campbell D, Reiter L, Picotti P, et al. ATAQS: A computational software tool for high throughput transition optimization and validation for selected reaction monitoring mass spectrometry. BMC Bioinformatics. 2011;12:78.
63. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, et al. Skyline: An open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics. 2010;26:966–8.
64. Martin DB, Holzman T, May D, Peterson A, Eastham A, Eng J, et al. MRMer, an interactive open source and cross-platform system for data extraction and visualization of multiple reaction monitoring experiments. Mol. Cell. Proteomics. 2008;7:2270–8.
65. Prakash A, Tomazela DM, Frewen B, MacLean B, Merrihew G, Peterman S, et al. Expediting the development of targeted SRM assays: Using data from shotgun proteomics to automate method development. J. Proteome Res. 2009;8:2733–9.
66. Randall SA, McKay MJ, Molloy MP. Evaluation of blood collection tubes using selected reaction monitoring MS: Implications for proteomic biomarker studies. Proteomics. 2010;10:2050–6.
67. Chang C-Y, Picotti P, Hüttenhain R, Heinzelmann-Schwarz V, Jovanovic M, Aebersold R, et al. Protein significance analysis in selected reaction monitoring (SRM) measurements. Mol. Cell. Proteomics. 2012;11:M111.014662.
68. Pepe MS, Zheng Y, Jin Y, Huang Y, Parikh CR, Levy WC. Evaluating the ROC performance of markers for future events. Lifetime Data Anal. 2008;14:86–113.
69. Cortes C, Vapnik V. Support-Vector Networks. Mach. Learn. 1995;20:273–97.
70. Bennett KP, Campbell C. Support vector machines: hype or hallelujah? ACM SIGKDD Explor. Newsl. 2000;2:1–13.
71. Karatzoglou A, Meyer D, Hornik K. Support Vector Machines in R. J. Stat. Softw. 2006;15:28.
72. Chang C-C, Lin C-J. LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. 2011;2:27:1-27:27.
73. Huber L. Validation and qualification in analytical laboratories. Inf. Healthc. , London,UK. 2007.

74. EudaraLex. Annex 15: Qualification and Validation. EU Guidel. GMP Med. Prod. Hum. Vetterinar USE. 2015;4:1–16.
75. ICH. ICH Topic Q2 (R1) Validation of Analytical Procedures : Text and Methodology. Int. Conf. Harmon. 2005. p. 17.
76. Chatziharalambous D, Lygirou V, Latosinska A, Stravodimos K, Vlahou A, Jankowski V, et al. Analytical performance of ELISA assays in Urine: One more bottleneck towards biomarker validation and clinical implementation. PLoS One. 2016;11.
77. Nakamura K, Kasraeian A, Iczkowski K a, Chang M, Pendleton J, Anai S, et al. Utility of serial urinary cytology in the initial evaluation of the patient with microscopic hematuria. BMC Urol. 2009;9:12.
78. Kumar A, Kumar R, Gupta NP. Comparison of NMP22 BladderChek test and urine cytology for the detection of recurrent bladder cancer. Jpn. J. Clin. Oncol. 2006;36:172–5.
79. Mahnert B, Tauber S, Kriegmair M, Nagel D, Holdenrieder S, Hofmann K, et al. Measurements of complement factor H-related protein (BTA-TRAK assay) and nuclear matrix protein (NMP22 assay)--useful diagnostic tools in the diagnosis of urinary bladder cancer? Clin. Chem. Lab. Med. 2003;41:104–10.
80. Huber S, Schwentner C, Taeger D, Pesch B, Nasterlack M, Leng G, et al. Nuclear matrix protein-22: A prospective evaluation in a population at risk for bladder cancer. Results from the UroScreen study. BJU Int. 2012;110:699–708.
81. Filip S, Vougas K, Zoidakis J, Latosinska A, Mullen W, Spasovski G, et al. Comparison of Depletion Strategies for the Enrichment of Low-Abundance Proteins in Urine. PLoS One. Public Library of Science; 2015;10:e0133773.
82. Granger J, Siddiqui J, Copeland S, Remick D. Albumin depletion of human plasma also removes low abundance proteins including the cytokines. Proteomics. 2005;5:4713–8.
83. Wood MW, Nordone SK, Vaden SL, Breitschwerdt EB. Assessment of urine solute and matrix effects on the performance of an enzyme-linked immunosorbent assay for measurement of interleukin-6 in dog urine. J. Vet. Diagn. Invest. 2011;23:316–20.
84. Trivedi D, Messing EM. Commentary: the role of cytologic analysis of voided urine in the work-up of asymptomatic microhematuria. BMC Urol. 2009;9:13.
85. Smeraglia J, Baldrey SF, Watson D. Matrix Effects and Selectivity Issues in LC-MS-MS. Chromatographia. 2002;55:95–9.
86. Song F. “Cross-talk” in scheduled multiple reaction monitoring caused by in-source fragmentation in herbicide screening with liquid chromatography electrospray tandem mass spectrometry. J. Agric. Food Chem. 2011;59:4361–4.
87. Zheng JJ, Lynch ED, Unger SE. Comparison of SPE and fast LC to eliminate mass spectrometric matrix effects from microsomal incubation products. J. Pharm. Biomed. Anal. 2002;28:279–85.
88. Rogatsky E, Stein D. Evaluation of matrix effect and chromatography efficiency: New parameters for validation of method development. J. Am. Soc. Mass Spectrom.

2005;16:1757–9.
89. Bastos P, Trindade F, Ferreira R, Leite-Moreira A, Falcão-Pires I, Manadas B, et al. EDTA-functionalized magnetic nanoparticles: A suitable platform for the analysis of low abundance urinary proteins. Talanta. 2017;170:81–8.